

THE MORAL SELF. AN EXPLORATORY STUDY OF MORAL AGENCY AND
CONCEPTIONS OF SELF.

BERİL SÖZMEN

BOGAZİÇİ UNIVERSITY
2011

THE MORAL SELF. AN EXPLORATORY STUDY OF MORAL AGENCY AND
CONCEPTIONS OF SELF

Thesis submitted to the
Institute for Graduate Studies in the Social Sciences
in partial fulfillment of the requirements for the degree of

Ph.D.

in

Philosophy

by
Beril Szmen

Boğaziçi University
2011

Thesis Abstract

Beril Sözmen, “The Moral Self. An Exploratory Study of Moral Agency and Conceptions of Self”

The main question directing this thesis is concerned with the implications of the nature of the self on the one hand and the conceptions that individuals have of their selves on the other for their moral judgement, their moral behaviour and their theory of morality. This aim is pursued along two main lines of enquiry. The first line is concerned with a conceptual clarification of the terminology used in the description and evaluation of moral agency as well as a broad overview of some of the most influential conceptions of self. The second line of enquiry is a critical exploration of the argument from a theory of relational self to a procedural moral theory. The results of this enquiry indicate that a relational theory of self as exemplified by Martin Buber and Arne Naess lays the foundation for a particularistic and dialogical moral theory, which focuses on the procedural aspects of moral agency and the role of authenticity therein.

Tez Özeti

Beril Sözmen, “Ahlaki Benlik. Benlik Anlayışlarının ve Ahlaki Eylemliliğin Keşifsel İncelemesi”

Bu tezin amacı, benliğin doğasının ve bireylerin benlik kavramlarının, onların ahlaki muhakeme, ahlaki eylem ve sahip oldukları ahlaki kuramlar üzerindeki etkisini sorgulamaktır. Bu amaç doğrultusunda iki temel yol izlenmektedir. İlk inceleme, ahlaki eyleyciliğin tanımlanmasında ve değerlendirilmesinde kullanılan terminolojinin kavramsal açıdan netleştirilmesi ve en yaygın benlik kuramlarına genel bir bakıştan oluşur. İkinci inceleme hattı ise, ilişkisel benlik anlayışlarına dayanan iki prosedürel ahlak kuramını eleştirel olarak ele alır. Bu iki yönlü incelemenin vardığı sonuç, Martin Buber ve Arne Naess örneklerinde işlenen ilişkisel benlik kuramlarının, ahlaki eyleyciliğin prosedürel yönlerine ve sahiciliğin rolüne odaklanan, partikülarist ve diyalojik bir ahlaki kuram için gerekli temellendirmeyi sağladığı yönündedir.

CURRICULUM VITAE

NAME OF AUTHOR: Beril Sözmen
PLACE OF BIRTH: Ankara, Turkey
DATE OF BIRTH: 02 December 1976

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

Boğaziçi University.
Birkbeck, University of London. January 2009-June 2006, visiting research student.
University of Cologne, June 2010-August 2010, visiting research student
University of Bonn.

DEGREES AWARDED:

Doctor of Philosophy in Philosophy, 2011, Boğaziçi University.
Master of Arts in Educational Sciences, 2003, Institute of Educational Sciences, University of Bonn.

AREAS OF SPECIAL INTEREST:

Moral Philosophy and Theories of Self

PROFESSIONAL EXPERIENCE:

Research Assistant, Department of Humanities and Social Sciences, Istanbul Technical University, since 2005

GRANTS AND SCHOLARSHIPS: TÜBİTAK International Research Fellowship Programme

PUBLICATIONS:

Sözmen, Beril. "Der Zusammenhang von sozialer Herkunft, Habitus und Bildungskarriere. Eine vergleichende Gegenüberstellung der Reproduktionstheorien von Siegfried Bernfeld und Pierre Bourdieu" Master's Thesis, University of Bonn, 2003

İdemen, Beril & Mert, Ayşem. "Peter Singer'le Söyleşi: Siyaset, Ahlak ve Gelecek" *Üç Ekoloji Doğa, Düşünce, Siyaset, Yeşil Politika ve Özgürlükçü Düşünce Seçkisi* 4 (2005)

İdemen, Beril. "Kim Korkar Pierre Bourdieu'den?" *Mesele* 4 (Nisan 2007)

İdemen, Beril. "Sosyal Köken, Habitus ve Eğitim, Pierre Bourdieu'nün Yeniden-Üretim Kuramı" In: Nurhan Yentürk, Yörük Kurtaran, Gülesin Nemutlu (Eds.). *Türkiye'de Gençlik Çalışması ve Politikaları*. İstanbul Bilgi Üniversitesi Yayınları, 2008

Sözmen İdemen, Beril. "Keeping a Sense of Self. Pathologies and Preferences of Self and Agency". *Phenomenology & Mind*, Phenomenology Lab, San Raffaele University Publishing House (forthcoming)

ACKNOWLEDGEMENTS

I am grateful to my advisor Assist. Prof. Karanfil Soyhun for her assistance during the course of this study. I also thank my committee members, Prof. Berna Kılınç, Prof. Stephen Voss, Assist. Prof. Lucas Thorpe and Assist. Prof. Aliye Kovanlıkaya for their support.

I thank TÜBİTAK for enabling me to pursue part of my research at Birkbeck College, University of London and Prof. Costas Douzinas for his kindness.

I am grateful to Dr. Wolfgang Krone, who awoke my interest in dialogical ethics when I first started studying philosophy, who invited me to do part of my research with him at the University of Cologne and who gave me much appreciated support in every way.

I am also grateful to Dr. Markus Rieger-Ladich, for his unvarying confidence in me as well as for his readiness to help me whenever help is needed.

I am indebted to Prof. İřtar Gözaydın in innumerable ways.

Finally, I thank my parents for providing me with everything that made this dissertation possible.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION.....	1
CHAPTER 2. MORAL AGENCY.....	8
The Moral Community.....	9
Moral Agents.....	12
The Moral and the Non-Moral.....	14
The Criteria of Moral Agency.....	20
Moral Patients.....	37
CHAPTER 3. CONCEPTIONS AND SENSES OF SELF.....	45
Conceptual Disambiguation.....	45
Kinds of Conceptions of Self.....	49
Character and No-Character Theories.....	54
Sense of Self.....	58
The Self and the Brain.....	66
The Shifty Self.....	69
Empathy.....	71
Localising the Self.....	74
CHAPTER 4. MARTIN BUBER’S MORAL ONTOLOGY.....	84
General Introduction.....	84
The Moral Situation.....	91
The I-It and the I-You.....	97
Self and Moral Agency in Buber.....	108
The Genesis of the I-It and the I-You.....	117
The <i>Wesen</i> of a Human Being.....	119

Dialogue and Monologue.....	130
The Affirmation of the Other.....	136
Authenticity.....	141
Philosophical Anthropology as Anthropocentrism.....	142
The Place of Divinity in Buber’s thought.....	145
Buber’s Moral Philosophy.....	149
On Good and Evil.....	155
Demands and Perfectionism.....	158
An Evaluation of Buber.....	162
CHAPTER 5. ARNE NAESS AND THE ECOLOGICAL SELF.....	187
Pluralism and Critique of Modernity.....	188
Ecosophy.....	190
Naess’ Conception of Self.....	191
The Link between Self and Morality.....	196
CHAPTER 6. CONCLUSION.....	199
Outlook.....	222
REFERENCES.....	227

CHAPTER 1

INTRODUCTION

Much of our lives rely on the idea that we are agents in general and moral agents in particular. We consider human beings to be capable of action, rather than only of behaviour and we consider them to be the source of action. The human moral agent as a theoretical construct is positioned in the cross-section between the realm of moral philosophy and conceptions of self. The main question directing this thesis is likewise concerned with both: What are the implications of the nature of the self on the one hand and the conceptions of individuals of their selves on the other for their particular brand of moral agency? As this question joins the moral agent in the cross-section between self and morality, an attempt at answering it must involve separate investigations into both areas as well as the evaluation of possible links between them.

In the case of the self, such an investigation must take into account the unique status of its subject matter as accessible from both a first person view and a third person one. It is this feature of the self or more generally of consciousness that has led to great divides in both philosophy and psychology. The behaviouristic school of thought has insisted that a scientifically rigorous approach must exclude the first person view and can legitimately only study the relations between externally observable stimulus and behaviour. Phenomenologists on the other hand focused on the first person view of consciousness. Philosophers and psychologists in recent decades have been trying to overcome this dichotomy. Owen Flanagan for example

calls for a unifying approach which he terms *the natural method*: “In making decisions about the nature and function of conscious mental states, or states with conscious components, consult the phenomenology as well as the psychological and neuroscientific research” (Flanagan 2009, p. 62). In the following, I shall follow Flanagan in taking into account both approaches without equating consciousness and its investigation with the self.

What I am concerned with is not the link between ontological theories on self, i.e., metaphysics of personal identity and ethics, but rather the link between philosophical anthropology or moral psychology and ethics. Do the challenges posed by sceptics of the self undermine ethics or does the possibility of understanding the self as a dynamic and inter-relational construct, based on more accurate moral psychology pave the way for an alternative approach to morality? In this complex of self, self-representation and ethics, my focal point will be moral agency; i.e., I shall study the various aspects of selfhood in so far as they contribute to the individual as a moral agent. I shall argue that strands of normative ethical theories, whose anthropological premise considers humans to be relational beings, such as the dialogical ethics espoused by Martin Buber and the ecosophy of Arne Naess are advantageously positioned to contribute to a comprehensive picture of the moral agent and her moral situatedness.¹ In addition, especially Buber’s ontology of

¹ I am borrowing the term ‘situatedness’ from cognitive science, where it is used to mean that the agents’ „behaviour and cognitive processes first and foremost are the outcome of a close coupling between agent and environment“ (Lindblom & Ziemke, 2002). In Artificial Intelligence, this theory led to the idea that cognitive development was so dependent on social situatedness that also artificial systems might need some sort of social interaction in order to develop (ibid). A related additional insight provided by studies into the development of human-like cognitive functions is the importance of embodiment for self and thereby agency, i.e., the necessity of a human-like body with all its senso-motoric ‘equipment’ (ibid). On a more ‘biological’ level, the importance of social contacts and impulses in the development of cognitive faculties has repeatedly been observed in cases of hospitalisation or sensory deprivation in experiments on animals or observations on children. While there is no doubt that deprivation of social interaction hinders the development of cognitive functions, the converse is by no means settled, i.e., it remains to be seen whether humanoid robots endowed with human-like bodies and social interaction will develop consciousness.

the In-between will open the possibility of adding the second-person approach to the investigation of the human as a moral agent.

The question of self is relevant to ethics in two ways; as an actor and as a deliberator. This is the case if we regard the self as the originator of moral judgements on the one hand and of moral actions on the other. The standard idea is that there is a straightforward causal relationship between attitudes, intentions and action. The agent behind the standard action theory is one that turns up in a number of guises. One of the familiar figures is that of the rational agent who masters his emotional cravings and learns to develop a stoical stance towards outer circumstances, aspiring to a state of ataraxia or autarchy. This figure has proved to be immensely resistant and its modern variant is "the ideal of the disengaged self, capable of objectifying not only the surrounding world but also his own emotions and inclinations, fears and compulsions, and achieving thereby a kind of distance and self-possession which allows him to act 'rationally'" (Taylor 1989, p. 21).

This idealised figure of the moral agent is often coupled with a perceived need for being single or unified. Socrates' argument for self-knowledge is a typical example both of this holistic tendency and of the awareness of it:

Those who gain this self-knowledge see that their happiness depends on psychological integration, or wholeness. We need to be at peace with ourselves. Inner conflict is a threat to happiness. Disharmony involves slavery to madness, and allows the beast in man to gain control (Glover 2001, p. 27).

Exemplary for the traditional conception of self and the link between self and (moral) agency is also a position described by Sorabji that "much of our ethics and agency depends on the same person – and that includes ourselves – being the owner of different activities and experiences, and on our recognizing this. The recognition involves a sense of a unitary self" (Sorabji 2006, p. 245).

Two senses of this tendency to unity or wholeness should be distinguished. The Socratic view mentioned above concerns synchronic identity; it is the urge to feel as a singular or unified being. A diachronic urge on the other hand is related to the idea of self, which “applies to individual humans and higher animals. Each of them needs to relate itself to the world in terms of *me* and *me again*” (Sorabji 2006, p. 4). If it is true that having a self involves having two such needs, what could be the reason for it? One can speculate that there might be an evolutionary force, which makes it more advantageous for the individual’s survival to have a diachronic sense of her self, since this would not only enable her to pursue projects which are necessary for her survival but also maintain a general concern for the well-being of her future selves, which would also increase her chances of survival.

Strawson argues against the idea that a sense of self is necessarily diachronic, i.e., that it involves an idea of the self as something that persists over time: “One can have a vivid sense of oneself as a mental self, and a strong natural tendency to think that that is what one most fundamentally is, while having little or no interest in or commitment to the idea that the I who is now thinking has any past or future” (Strawson 1999, p. 14). Conscious and active striving to overcome or undermine this sense of temporal unity and wholeness is part of other traditions, which emphasise the hindrance to life that is caused by adhering to a unified² sense of self. Depending on the tradition, what is being hindered can be understood as the good life, in the sense of *ars vivendi* or as the morally right life or a combination of the two. Buddhist ideals of the dissolution of the self or Schopenhauer's idealisation of quietism are

2 I distinguish between the terms *unified* and *single*: The term *unified* implies that there is a multitude of functions or parts, which either function in such a way as to produce a whole or which are experienced as such. A *single* self would be one that has no parts. The former view might be termed Aristotelian; the self has many parts because, human nature has metaphysical affinities to the nature of plants and animals. The platonic view has a unified aspect as well in that rationality is the grand unifier and leader of “lesser” parts of the self. The latter approach is nearer to the Cartesian one, in which the self is understood as pure ego, pure cogito which does not consist of any parts.

examples of attempts at transcending self by reducing desire while the same goal is envisaged with techniques bent on an aesthetic handling of desire itself.

Whether the self is considered as the unified and stable unit of (moral) judgement or action and as such necessary for any moral theory or whether it is considered to be a hindrance to developing a broader sense of being, which encapsulates and thereby postulates as moral units other beings than the moral ego, the idea of self and the idea of morality are deeply intertwined. Another example of this interdependence is the view that the normative stance of a subject, the sense of what is good is an integral part of the self. Taylor, for example, follows Iris Murdoch in arguing that it's not the subject or the self that is sovereign, as a utilitarian picture of agency would suggest but that "Selfhood and the good, or in another way selfhood and morality, turn out to be inextricably intertwined themes" (Taylor 1989, p. 3). This way of explicating moral agency is similar to the ancient conception, reproduced by the romantics of the role of the artist in the creation of an artwork. Just as the artist in the process is interpreted as the mouthpiece of Gods or of some ideal realm of art, similarly the moral good is channelled through the subject.

There is a long tradition of philosophical and/or theological approaches to the question of the self. Historically, the most common view is that of an immaterial, immortal soul, which can be made up of two or three parts, as the Platonic self or is unified as in Abrahamic religions. A second view, more common since modernity, is the conception of self as a "constantly changing process of interrelated psychological and physical elements, later phases of which are appropriately related to earlier phases" (Martin and Barresi 2003, p. 1). This view involves memory as an important aspect of self. Some contemporary views criticise such 'intrinsic' conceptions of self and argue that 'extrinsic' factors of personal identity – relations of the person to the

outside world, so to speak – must be taken into account. Other views argue that such a conception of identity doesn't matter in survival (Parfit 1986). Yet another view proposes to replace the idea of identity as psychological or biological continuity with momentarily existing time-slices of temporal beings. Since my interest is not in the unity of an individual over time but rather the unity of the self at a particular time and in relation to (moral) judgement and action, I shall not dwell on these latter theories.

I shall argue that this very quality, this urge for unity can lead agents to less than optimal capacity as a moral agent. If instead, it is possible to show that unity of self is not a *sine qua non* for (moral) agency, then we might be able to work on a relational, intersubjective morality which does not rely on saving face. Doing so, we might find ways to make the realisation of many a vice like self-deception or post-hoc rationalisations less necessary and thereby support many a virtue such as attentiveness or openness. I am going to argue that such a move is possible by understanding the moral self as a relational construct. When I use the term 'relational', I shall follow Harré in his description of the attributes of an individual's selfhood: "Not only are they defined in terms of relations to other people, and to the characteristics of those others reciprocally, but they are in constant flux and the relations to the social and material environment shift and change" (Harré 1998, p. 7). While developing such an understanding of the self, I shall adapt Leibniz' monads to the basic units of a multiple and dynamic self: "Individuals are not substantial entities, but aggregates of monads, and thus complex assemblages of separate determinations" (Seigel 2005, p. 78).

I shall then go on to consider Martin Buber's dialogical philosophy, which distinguishes a momentary state of self as the effect of a relation between the *I* and

either *You* or *It*. I shall also discuss another philosopher who has developed a relational model of self, Arne Naess. Unfortunately Naess has written little on his conception of the relational and – in his case – the ecological self, so that there is a quantitative imbalance between the explications of the two thinkers. Nevertheless, I take Naess' position to be just as important as Buber's, especially considering the implications that his discussion of the moral self has for applied ethics.

I shall conclude by arguing that there is a case to be made for a moral theory, which focuses on moral situations and on the processes, by which moral agents encounter one another and come to make morally relevant decisions. Combining a procedural approach to morality with the view that a relational conception of self is the best position at hand, I shall conclude by arguing that authenticity deserves a pivotal role in such a procedural moral philosophy.

CHAPTER 2

MORAL AGENCY

Moral agency is a single term for a whole bundle of inter-related questions, some of which are among the central issues of philosophical enquiry. To attempt a conceptual analysis of the term is an enterprise in its own right. Therefore, in the following, I shall rather focus on the implications that various basic conceptualisations of moral agency have for related question concerning the boundaries of the moral community. The emphasis will be not so much on an analysis of moral agency per se but rather on a practically oriented exploration of moral entities, their relations and their implications for a moral theory.

An enquiry into moral agency and moral patiency; in other words, the question concerning the criteria of being a member of the moral community belongs to the subject matter of a number of areas of philosophy, as well as many other disciplines. In philosophy, both ethics and the philosophy of mind are directly related to the establishment of the conditions of agency and self, while political philosophy relies heavily on varying anthropological views and normative theories about the qualities, capabilities and needs of the citizenry. Comparably, related disciplines like psychology, sociology, educational sciences and anthropology study the psychological, cognitive and social aspects of agency. A comprehensive study of the issue must therefore involve an inter-disciplinary approach, but within the framework of this thesis, I can do no more than refer to approaches in other disciplines in so far as they are relevant to my overall aim.

I shall explicitly exclude one question, which is in fact central to the enquiry. The question concerning free will is one that must be answered affirmatively for any postulation of moral agency to make sense. This is not necessarily the case for the discussion of moral patiency; the question whether a being is to be counted as a moral patient does not necessarily demand that the being be a moral agent – or so I shall argue. However, considering moral agency without presupposing that the agent is endowed with free will makes little sense. In this paper, I shall assume that, unaccountable as it may seem, free will exists and that it is possible for human beings to be moral agents in this strict sense of freedom as opposed to causal determinism. Whether free will can also be understood as lack of coercion by psychological or social forces, as the compatibilist approach will have it, is quite another question, the affirmation or negation of which is also central to the enquiry into moral agency. Nevertheless, I shall keep the question concerning mental causation in brackets and assume that at least some beings may be regarded as having free will so that my quasi-anthropological assumption will involve moral agents as the initiators of action. Socio-political freedom from coercion as a distinct question on the other hand, will make up one of the possible criteria of the development of moral character without being assumed outright. In short, I shall assume that determinism is false but that there are sociological and psychological determinants of moral behaviour.

The Moral Community

In the following, I shall first attempt a rough conceptual clarification of what it entails to claim that a being is a moral entity: Which criteria are necessary for a being

to be regarded as a member of the moral community? Who is the "we" one is morally bound to consider while thinking and acting? The methods that approach this question seem to understand it as a question of the correct application of sets encompassing moral agents, moral patients and non-moral beings. How one must act and how one must be treated thus becomes a question, which is directly linked to one's "nature", which – if postulating species-specific natures' is not an inadmissible essentialist move – is an empirical question. If for example, the determining quality of moral patiency is sentience, then what is needed is not only a philosophical task of conceptual analysis of 'pain' or a phenomenological analysis of pain, but also a biological one.

The philosophical question of 'belonging' in the more narrow sense involves a set of questions about justice and rights, starting with the Stoic concept of *oikeiosis* and its concentric circles of belonging (Sorabji 1993, p. 122 ff). *Oikeiosis* as the opposite of alienation is the sense of appropriating beings as belonging to one's self. What makes up my self and my moral concern can begin with my appropriation of my body (famously called into question by Descartes) and can end with the whole set of (sentient) beings. The stoic concept of *oikeiosis* takes the moral community to be that of rational beings; an example of reciprocal ethics, which is challenged by Bentham.

An enquiry into the figures of the moral sphere is from its very beginning loaded with various assumptions, which are in themselves anything but unproblematic. To begin with, one assumes that there is a distinct sphere of morality; one that is not necessarily over and above but necessarily distinct from the sphere of material entities. However, it is important to maintain the distinction between materialistic and idealistic ontological positions on the one hand, and the distinction

between a sphere, in which so-called laws of nature reign and one that is the realm of freedom on the other. Possibly the most prominent of these latter type of distinctions is the one made by Kant when he postulates a kingdom³ of ends (Kant 1999, p. 59). As metaphysically obscure as such a kingdom may initially sound, Kant expressively defines it as “the systematic connection of various rational creatures via common laws” (Kant 1999, p. 60). In the following, I, like Kant, shall bypass considerations about the possible ontological nature of the moral sphere and shall understand it as a dynamic and systematic constellation of relations among moral entities. Therefore, I shall give priority to an enquiry concerning the nature of moral entities, the results of which can provide a possible outlook for an ontological enquiry of the moral sphere for anyone inclined to attempt it. In such an approach I shall keep in brackets not only metaphysical considerations but also meta-ethical ones as to the nature of ethical concepts and their relations.

The enquiry thus turning to moral entities, another and final exclusion seems appropriate. The term 'entity' is just as loaded and problematic as that of 'sphere' or 'realm' but since the very aim of this enquiry is to determine the necessary conditions of belonging into the moral community, I shall use the term as denoting moral beings, the qualifications of which are to be specified in due course. I shall begin with considering the moral agent, who is traditionally given more weight than the moral patient. While considering the criteria that entitle a being to be regarded as a moral agent and subsequently dwelling on moral patients, it will gradually appear

3 The original German term *Reich* has traditionally been translated into English as 'kingdom', which emphasizes the region in which a power reigns, whereas the word *Reich* could arguably better translated as 'reign', which carries both the idea of the – special and temporal - sphere and more importantly the characteristic dominant quality which turns this sphere into what it is – in this case one of autonomous morality.

that there is reason to argue for the existence of a gap between agent based normative theories as opposed to patient based ones.

Moral Agents

The term 'agent' is complicated in that it seems to intersect with or to be used more or less synonymously with a number of other terms such as 'subject', 'person', 'self' etc. When I use the term 'agent', I shall be leaning more towards the aspect of (moral) action; again quite conservatively: "An 'agent' is someone who is contemplating an action, has already acted or is presently acting" (Bartky 2002, p. 31). I shall use the term 'self' on the other hand with a broader meaning, which encompasses also passive, purely perceptive mental events. Finally there is the broadest sense of self as a human agent, which includes "the sense of inwardness, freedom, individuality, and being embedded in nature" (Taylor 1989, ix) as a portrayal specific to Western thought. Whether or not 'self' is a meaningful concept, that there is a sense of self in many people is indubitable. The question concerning this sense of self is a phenomenological one, whereas a conception of self might be understood as a more conceptual, 3rd person view of self. Strawson introduces a further distinction along with the term 'conceptual phenomenology', i.e., the distinction between "the conceptual structure of the sense of the self, the structure of the sense of the self considered (as far as possible) independently of any emotional aspects that it may have" (Strawson 1999, p. 2).

These attempts at conceptual clarification should not be understood as ultimate definitions of the terms 'self' and 'agency' but rather as working definitions, which should enable us to turn to the more interesting questions concerning them:

What are the conditions that an agent has to fulfil so as to be rightfully called a moral agent; i.e., what are the personal characteristics that one needs to possess? This question dwells on the characteristics of a moral agent and involves trivial ones (being alive, being conscious etc.) as well as more contentious ones (rationality, consistency etc.). What sort of a type must an entity be in order to qualify for moral agency? Can only individuals be moral agents or also groups, corporations or even nations? Can collectives be regarded as morally obligated to perform or omit certain acts? Can they be called to responsibility, found blameworthy or punished collectively?

On closer look, the question regarding moral agency breaks up into at least three distinct but closely related questions. The first question concerns the conditions that an agent has to fulfil so as to be rightfully called a moral agent; i.e., what are the personal characteristics that one needs to possess? Secondly, what are the characteristics of the actions, which are to count as conceptually relevant for a moral action? Finally, what sort of a type must an entity be in order to qualify for moral agency? While the first question dwells on the characteristics of a moral agent, the last question aims to investigate the conditions to count as a moral agent in the first place: i.e., can only individuals be moral agents or also groups, corporations or even nations? Can collectives be regarded as morally obligated to perform or omit certain acts? Can they be called to responsibility, found blameworthy or punished collectively?

Needles to say, these questions are interrelated. The quest for a comprehensive list of the personal characteristics that allow someone to be called a moral agent will probably have to be met also in the case of collectives. However, the question whether these characteristics are inextricably bound to human

individuals is a distinct one. If the answer to the latter question turns out to be affirmative, then the problem of collective action becomes even more intractable than it already is. For how can certain collectives, e.g., a generation be regarded as morally praiseworthy or culpable for an action, when, for instance, an adult individual bonobo allegedly can't be? Similarly, the qualities of actions, (i.e., whether they are habits or rational choices etc.), which have to be met in order to speak of moral agency, are related with anthropological and zoological questions about human and animal agency on the one hand and sociological questions about collective action on the other.

The moral and the Non-moral

This part will deal specifically with the question of which elements of an action make it a distinctively moral one. I shall try to answer this question by trying to locate differences between moral and non-moral actions. I shall argue that among the answers given, Mill's Harm Principle is the most promising but not unproblematic.

Some meta-ethical remarks are in order at this point. The conflict between moral realists and subjectivists turns on the question whether it makes sense to speak of right and wrong moral stances as corresponding to objective state of affairs in the world. In this context I shall bypass the meta-ethical question concerning objective moral truths and opt for the less ambitious stance of a more or less. Although this stance is not sceptic-proof, it is not as open to the relativistic or subjectivistic criticism as absolutistic moral theories. I shall assume⁴ that moral realism is correct,

⁴ I use the term *assumption* in the foundational sense that Owen Flanagan defines: "An assumption is foundational or metaphysical if it articulates, without defense, what is taken to be a settled matter of philosophy – e.g. that persons exist; that there are multifarious character traits, many of which subserve moral life and can be used to predict and explain behavior; that some actions are voluntary,

although there are good reasons to be sceptical about the existence of a universe of moral truths. It is possible to speak about better or worse in moral judgement and moral action, without having to postulate ‘queer’ facts about a realm of universal moral truths. Moral judgements and arguments can be more or less impartial, consistent, complete, critical and thereby more or less cogent.

In recent decades there has been a return to a broader view of morality, which does not exclude self-regarding aspects of the good life and which views it as permissible “strong evaluation” to approve or disapprove of how a life fares even if it is morally impeccable (Taylor 1989, p. 4). This is the question regarding the criteria for the life well lived in a more comprehensive sense than solely the moral in the narrow sense of right action. What underlies the view of the good life is the belief that there are certain moral reactions, which are “instinctive, physical, and animal, not unlike vomiting with disgust at stinking and rotting things, fainting with fear of falling, and such like” (Kerr 2004, p. 93). Although such reactions can be trained to be more sensitive or conversely to be overcome, Kerr argues that there is a universal core of human moral reactions, which are not part of a utilitarian calculation of the right, but a deep-set belief in the good (ibid.). The historical turn has been one from such a holistic view, which encompasses the good to one that deals solely with what is right to do, partly because basing the moral life on the Good requires moral realism, a transcendental realm of moral truths, which is unpalatable to many modern thinkers: “To understand our moral world we have to see not only what ideas and pictures underlie our sense of respect for others but also those which underpin our notions of a full life” (Taylor 1989, p. 14).

some are not; that responsibility tracks voluntariness; and so on” (Flanagan 2009, p. 66).

What are the characteristics of an action if it is to count as moral? This question is not concerned with the distinction between moral and non-moral actions but rather with the qualities of events; i.e., whether they are habits, rational choices, etc and includes the question of whether collective events can be understood as actions. As soon as one sets out to investigate the necessary conditions for moral agency one stumbles across another implicit assumption; namely the claim that agency is one thing and moral agency something else. The qualification ‘moral’ implies that there are other sorts of agencies or even agency *sic et simpliciter*. It is therefore important to denote clearly what is meant by moral as opposed to other forms of agency.

It is customary to distinguish the realm of morality from others by referring to one or more specifically moral criteria, which must be fulfilled. Various contestants have been brought forward in the history of normative moral philosophy as opposed to purely descriptive accounts akin to those of anthropology or sociology on the one hand or those with a historical and deconstructive agenda as the Nietzschean attempt on the other. Rationality and accountability is a pair often encountered.

Universalisability (Kant) is another criterion put forward as a distinguishing mark of moral statements, while the harm-principle (Mill) is a further one. With the gradual emancipation from moral codes based on what was believed to be divine command, the emphasis turned on the distinguishing features of morality as opposed to other realms, particularly the aesthetic. With the advent of classical liberal thought, it became widely acknowledged that how an individual lives her life is her own concern unless – and it is this very attempt at demarcating this „unless“, which came to be regarded as the boundary between individual liberty on the one hand and morality and law on the other: “As soon as any part of a person's conduct affects

prejudicially the interests of others, society has jurisdiction over it, and the question whether the general welfare will or will not be promoted by interfering with it, becomes open to discussion. But there is no room for entertaining any such question when a person's conduct affects the interests of no persons besides himself" (Mill 2002, p. 78).

As straightforward as this demarcation seems to be, one closer look makes it become less so rather quickly. On the one hand, there is an ancient tradition of those who claim that morality is about how one ought to act in order to live a good life. This line of thought considers not only how our actions might affect others but what sort of a person one has to become in order to be able to claim – retrospectively from one's deathbed, so to speak – that one has led a good life. Taken to the extreme this approach might make the claim that the moral life is reducible to the aesthetic of existence; the focus is not on the recipient end of action but by the agent himself. The opposite extreme will find us claiming that there is not a single action, which can be regarded as purely aesthetical; nothing is imaginable, which does not affect other moral entities in one way or the other and therefore a sphere of individual liberty cannot exist. The case for this side is even more easily argued for: Can there really be an action, which has no effect on anybody else than the agent? The popularity of the slogan "My freedom ends where yours begins" notwithstanding, can there be any action, which the agent can perform without being morally bound to consider other, potentially involved moral patients? With the increasing sophistication of environmental ethics, feminist ethics or the animal liberation movement, what was considered to be purely personal and private choices of consumption, life-styles or employment are shown to have indirect or long-term impacts on common resources or other beings.

Nevertheless, the temptation to distinguish moral actions from non-moral ones is strong. Charles Taylor, for instance, distinguishes between weak and strong evaluation of actions or lifestyles, where strong evaluations are based on the “qualitative worth of different desires” (Taylor 1985, p. 16) while weak evaluations are concerned with decisions in life, in which the distinctions at hand are not such of worth. Whether one decides to eat muffins or bread and butter for breakfast is for Taylor a decision which involves weak evaluations of the two alternatives; the decision is based on nothing more than the whim of the moment and not on an evaluation of the worth of the two alternatives. While a weak evaluation contains nothing but a preference or desire for the path chosen, the Good which is sought in a strong evaluation is of a higher order. Therefore, decisions based on weak evaluations are concerned with contingent incompatibilities regarding circumstances, whereas decisions involving strong evaluations are incompatible in a more substantial sense: If the difference between being a brave person or a coward is a distinction of worth for me, then to succumb to the temptation to flee from a situation where bravery is required is substantially incompatible with acting courageously (Taylor 1985, p. 19). In this way, the question turns to one about self; there are two selves involved in these two approaches to life, one calculating one's desires and weighing between the expected amount of pleasure of each course, the other concerned with “deeper” issues about what is worthy of being chosen to live a good life (Taylor 1985, p. 23)

This method of drawing the boundary between the moral and non-moral is familiar enough. It postulates one set of actions, which are based on desires or preferences where it doesn't make a difference what one chooses, where one is free to act as one wishes, while the other set is distinctly normative in a moral sense.

However familiar this view might seem, it is clear that choices of holiday resorts or those between muffins or bread and butter are seldom as morally irrelevant as one would wish. Many social movements of the past decades focused on food choices, arguing that what one eats matters both politically and morally, while the damage caused by touristic choices is an environmental threat worldwide. It seems that the claim that choices and actions can be divided into two realms, a private and a public, a moral and a non-moral is problematic, without there being need to resort to the postulation of moral duties towards oneself.

Taking duties towards oneself into account complicates the case even further; such actions as seem to concern solely ourselves can be twisted into moral ones very easily. A virtue ethicist for instance would argue that from a certain age onwards a virtuous person is responsible for her own moral education and development so that even those actions which do not affect other people are morally relevant since they contribute to the development of one's moral character. To choose muffins over bread and butter could therefore be reprimanded as being self-indulgent, unhealthy or luxurious. A deontologist would speak of moral duties towards oneself in a similar vein and utilitarians would not be hard pressed to find examples for cases, in which self-control and self-denial increase utility in the long run. Nevertheless, it seems just as clear that these difficulties must not necessarily leave us in a position unable to make differentiations and distinctions at all. Whether one chooses to eat meat or not, for example, is a clear-cut case of moral action while one's sexual orientation or preferences – among “consenting adults” – is as non-moral as an action can get. Such clear-cut cases are exceptional however and the attempt to distinguish moral agency from agency *sic et simpliciter* is as difficult as it is necessary. I shall refrain from embarking on such a project in more detail in this context. Keeping in mind the

problems involved we can now proceed to clarify the characteristics of agents and their actions.

The Criteria of Moral Agency

At this point, yet another exclusion seems in order. The ontology of acts and events is a further necessary course of enquiry into a comprehensive study of moral agency and the moral community. A first ontological look at acts casts these as a class of events bound to specific agents; an act is an “‘exercise of agency' [...], an event with an 'agent'” (Thomson 1977, p. 24) and as such go beyond “‘happenings’”, the prototypical example of which would be falling asleep (ibid..). Yet, however the understanding of actions as events might be contested; I shall work with a naive understanding of action as a set of events caused by the practice or omission of bodily movements of agents based on their beliefs and desires (Hornsby 2004, p. 2). In the following, I shall consider the criteria of moral agency while disregarding ontological problems about acts and events.

One seemingly rather obvious criterion of being an agent is being alive. Whatever might be the answer to the question which living beings are eligible for agency, that one must be an animate being seems closed to doubt. Even if a bullet might be regarded as the object responsible for someone's death in the sense that it was the cause of the event, it is clear that the term 'agent' is reserved for animate beings. However, with the progress in cognitive sciences and artificial intelligence, the question now turns on what “animate” is supposed to mean. This line of enquiry quickly leads to the question of which cognitive elements of mind are necessary for agency and whether these can be artificially produced in certain beings, which are

not animate in the traditional sense. The latter question is irrelevant in this context and I shall not dwell upon it here. On the other hand the question what the necessary cognitive elements for conditions for agency are, is a central one, as is the supplementary question, whether these elements are sufficient.

Possible answers to the first question above seem far less problematic than those to the second. For a start, consciousness is an obvious candidate for a necessary condition of agency. I use the term consciousness here in the sense of being awake rather than asleep or in a coma. Whatever events may be caused by a comatose patient or a sleepwalker, consciousness in the sense of being capable of – at least – perceiving and commenting on one's actions must clearly be granted an agent. This cognitive capacity we can call consciousness in this most basic, quasi-biological sense.

A further set of postulated cognitive capacities is more complex: To be eligible for agency, a being must not only be conscious of one's actions, i.e., possess knowledge that one is acting in such and such a way but also have awareness about the wider context of the action, its reasons, possible consequences or the intentions and motivations of others and of the agent herself. The concept of awareness as such touches our intuitive ideas of morality, since it denotes a clear idea of what one is doing; an insight going beyond the causes of an action and including reasons.

Liminal cases between consciousness and awareness are found often in psychoanalytical literature; a primitive example of pop-psychoanalysis could be the urge a paedophile feels towards having sexual relations with children. She might act on the urge without having a clear notion of what she is doing and especially why she is doing it, while psychoanalytical awareness might provide her with insight into her personal biography and reveal the causes of her desire, e.g., a history of sexual

abuse. Awareness, in this example, denotes a more thorough knowledge into the causes and wider social and historical context of one's actions and one's self. Such awareness of one's self is not simply cognition but also active in the constitution of selfhood, the conception of self, i.e., the theory that one has of one's self might be argued to contribute to its being and nature.

The factors "animation", consciousness and awareness combined are still not sufficient for fulfilling a further condition of agency: Accountability or justification in the moral sense. Our paedophile might find herself enlightened about the genesis of her condition without this insight providing a justification for her behaviour. In this example, if the paedophile uses her newly gained knowledge about the causes of her behaviour as reasons for it, she would commit a genetic fallacy in the moral sense, confusing the origin of her behaviour with its justification. Accountability, as moral agency requires it, denotes more than being conscious of what one is doing or being aware of the intricate and hidden reasons for it; it requires that the given reasons are given as a justification of the action, which the agent can argue for their being right, reasonable and acceptable. Understood like this, accountability transports us directly into the heart of moral reasoning, independently of what strand of moral theory is used and whether the attempt at justification works in the chosen framework.

Part of a wider awareness of what one is doing, is a further criterion akin to the Aristotelian virtue of prudence (*phronesis*). To be a moral agent, one needs to have a clear notion of the relevant facts along with practical wisdom acquired with experience. Prudence, as a virtue, is "a truth-attaining rational quality, concerned with action in relation to the things that are good for human beings" (Aristotle, *Nicomachean Ethics* 1140b19). Prudence requires not only such cognitive skills that

are needed for theoretical or scientific reasoning but additionally experience in dealing with particulars and as such it is the sort of wisdom acquired with age. Thus, prudence is a virtue by its own right but requires other virtues to be acquired in order to function so that this condition of agency had best be understood as a compound of virtues. These must include not only the cognitive skills themselves but also the readiness to acquire them.

However, there is a catch in the postulation of such a meta-virtue, i.e., a virtue the object of which is to become virtuous. As such, this difficulty is part of the general difficulties regarding moral motivation. The will in a young person to acquire something, which, by its very nature is as yet closed to her and which can only be acquired in a long time and with constant practice and vigilance is in itself quite a remarkable virtue.⁵ What possible motivation a young person – or any person – could have for acquiring prudence in this narrow sense, for becoming virtuous in a wider sense or for being moral in the widest sense, remains quite obscure. Virtue ethics might postulate that an older person who has spent her time effectively in pursuit of becoming virtuous will thereby have acquired a sense of the necessity or the value of living virtuously; this will be part of both her theoretical and her practical wisdom. But if youth or inexperience excludes a person from having the latter, how can she be expected to be motivated to work towards acquiring it?

Apart from this point related to moral motivation, moral development and moral education however, it seems clear that the tradition of agent-based Aristotelian ethics reveals an important insight about the necessity of creating a moral character,

⁵ Kant makes a similar point: “There are certain moral endowments such that anyone lacking them could have no duty to acquire them. They are *moral feeling, conscience, love of one’s neighbour, and respect for oneself (self-esteem)*. There is no obligation to have these because they lie at the basis of morality, as *subjective* conditions of receptiveness to the concept of duty, not as objective conditions of morality. [...] To have these predispositions cannot be considered a duty; rather, everyman has them, and it is by virtue of them that he can be put under obligation” (Kant 1991, p. 200).

which not only includes the will to act morally right in a given situation but also accommodates the necessity to acquire the required cognitive and volitional skills, e.g., prudence. Without a stretch of the imagination, also deontologist or consequentialist ethics can provide such a perspective of moral agency; a deontologist for instance can argue for a wider sense of duties towards oneself which incorporate moral (self-) education, while rule utilitarians wouldn't find it too difficult to show that to work towards the establishment of moral character, be it in oneself or in the next generation, pays off in the long run.⁶ One could even argue that the whole judiciary system assumes such a didactic role and that virtue ethics does not have the monopoly on the subject of moral development; however, the discussion of moral agency would lose much if it lacked the study of moral character, its development and education.

Prudence or practical wisdom is a virtue necessary for two of the skills a moral agent must further master: She must have a clear notion of the relevant facts of the situation; i.e., she must be able to scan the available information, select the relevant ones, weigh them against each other and have a good memory to evaluate a particular situation in the light of past experiences. Additionally, she must have a clear notion of the consequences of her intended action in the particular context of the situation she finds herself in. As yet, these skills do not point to a specific conception of morality or anthropology. To be capable of judging the situation appropriately or to be able to estimate more or less accurately the likely consequences of her actions, an agent must neither be committed to one specific normative moral theory nor be a *homo economicus*, rationally calculating her self-interest in the transactions she is involved in. Nevertheless, without being in the

⁶ Thus Mill speculates: "Utilitarianism, therefore, could only attain its end by the general cultivation of nobleness of character, even if each individual were only benefited by the nobleness of others, and his own" (Mill 2002, p. 244).

position to correctly judge her situation, her options and the intended results, a being cannot be considered as a moral agent; whatever the moral tendencies of her considerations might turn out to be.

Related is a further criterion for moral agency, which is, at the same time among those most commonly postulated: Reason. A moral agent is – as practically all moral philosophers are unanimous in claiming – necessarily a rational creature; her decisions are based not on arbitrary and unreliable affective states but on consistent dictates of her reason. It is allegedly this rationality of humans that sets them apart from other animals. Related to rationality, consistency is a further criterion for moral agency; one, indeed, which is often used to argue against one's opponent. The set of one's moral judgements, as it is argued generally, must show internal consistency, either in the compatibility of individual judgements with each other, or with the actions they allegedly lead to. Rationality and consistency are among the most established criteria of moral agency. Even if the holistic approach of virtue ethics often criticises the insufficiency of rationality and pleads for a more thorough inclusion of the emotional aspects of agency, rationality – and thereby accountability – is a major aspect of moral agency.

Both the existence and the nature of affective components of moral agency themselves used to be strongly contested. Emotions are treated in a rather stepmotherly fashion in Kantian approaches to morality; their motivational impact on an agent must be nil in order for an action to be regarded as morally right. A notable exception is the feeling of respect. For Kant himself it is reason alone that can determine morality and not feeling of any sort, be it the wish to attain happiness, the urge to avoid pain or an imaginable moral feeling immediately related to the consciousness of having done good or bad: “Conscience” or “moral feeling” are

concepts that Kant does not regard as relevant to the moral worth of an action (Kant 1974, A 67). However, respect for the moral law is itself a feeling even though a peculiar one: It is based on an intellectual fundament and it is the only feeling, which is necessary for moral action and not contingent and subjective (Kant 1974, A 130).

The virtue ethicist's criticism of the incompleteness of a purely rational account of moral agency and the doubts of more contemporary thinkers as to the role of rationality in decision-making suggest that the Kantian understanding of agency leaves some important moral intuitions unsatisfied. This is demonstrated most clearly if we envisage a being which derives its moral motivation solely from judgements it rationally forms; Mr. Spock being a good example (Baggini 1997, p. 5). Mr. Spock, the fictional character from Star Trek, is a being devoid of an affective life. This, far from making him callous or selfish, makes him morally upright in the Kantian sense: He chooses his actions according to their rational value and behaves in the right way, without having a sense of empathy or sympathy with those who feel the consequences of his actions. Kant might have been impressed but how are we to judge such a creature as a moral agent? Mr Spock, it seems, condemned moral acts only when and because they were irrational – which perfectly suits the demands of the categorical imperative. And yet, there seem to be perfectly common and ordinary cases where the lack of the affective component of an action renders it worthless. Among these are such acts of consideration, friendship, care or love among friends, lovers or family members. Surely, one could easily dispense with or even spurn an act of kindness, which is done not out of genuinely felt consideration and interest in the others well-being but out of an impersonal sense of duty.

While the case of Mr. Spock is hypothetical, some psychopathologies provide real-life cases of aberrant moral characteristics. Psychopaths and autistic people, if

they can be taken as examples of moral failures⁷, are relevant for moral philosophical attention precisely because they seem to demonstrate cognitive skills, which can be normal or above average as compared with the general population but lack affective skills. Psychopaths seem capable of attributing mental states to others but they remain indifferent to the suffering of others, fail to recognise facial expressions of emotion (McGeer 2008, p. 230) and they have an incorrect and contradictory use of evaluative terms (Kennett & Fine 2008, p. 176). While such cases suggest that the affective component of moral life should be taken into account, the specific comparison of psychopaths and autistic people suggest that the role of empathy might have been overemphasised as cruel or criminal behaviour is more common among psychopaths, who are capable of correctly attributing mental states to others than among autistic people, who retain moral concern (McGeer 2008, p. 233). High-functioning autistic people seem to tend towards Kantian moral philosophy rather than Humean; their lack of affective skills lead them to concentrate on reason as a source of moral judgement (McGeer 2008, p. 234).

The example of Mr. Spock and some psychopathologies makes clear that it is right to give more weight to the affective components of moral agency; on the capacity of feeling empathy and sympathy with other sentient beings and on the feelings of remorse and regret about moral failings or even about moral tragedies the actors are not responsible for. This is an important point, not only because it provides a more comprehensive conception of moral agency, but also because it points to a conceptual gap in moral theory. If we understand moral philosophy exclusively as the question concerning morally right actions, then a world inhabited by creatures

⁷ As so often, for every sensational case of a psychopath or autistic person who does not demonstrate a moral sense, there is a counter-example. McGeer reports of an autistic young man “with perfect pitch and a passion for pianos”, whose own delight in piano music led him to argue that “there should be a constitutional amendment requiring every home to have a well-tuned piano” (McGeer 2008, p. 232).

like Mr. Spock would provide a great number of morally right actions, perhaps even the greatest possible number. However, the fact that one recoils from such an idea indicates the possibility that this hypothetical case might provide a counterexample as to the desirability of such a world and shows that the qualities of moral agents have moral value in themselves, even if the greatest number of morally right acts can be guaranteed by some other arrangement. Utilitarianism, as a prototypical case of an action-based moral theory faces the criticism that the agent is reduced to a vehicle contributing to this or that state of affairs. The gap, it seems, is between the moral evaluations of an agent on the one hand and of the act in the other; to have potentiality and actuality in realising right moral acts seems to be a necessary but insufficient criterion to be a moral agent.

This insight is shared by virtue ethics and care ethics. To demonstrate, one can imagine being the subject of two identical acts involving the realisation of one's personal interests. While a rule-oriented moral theory based on the equal rights of all moral patients to certain rights and more generally on justice would nominally provide both subjects with what they need, there are cases where the subjects would nevertheless choose one agent over the other as more morally praiseworthy and these seem to be the very cases that care ethics argues one should take into account more often. The character of the moral agent matters for the evaluation of the act; it is not enough that the agent acts rightly, but he must act for the right reasons and with the right affective and motivational state. As Aristotelian ethics claims:

Acts done in conformity with the virtues are not done justly or temperately if they themselves are of a certain sort, but only if the agent also is in a certain state of mind when he does them: first he must act with knowledge; secondly he must deliberately choose the act, and choose it for its own sake; and thirdly the act must spring from a fixed and permanent disposition of character (Aristotele 1996, p. 38)

All other things being equal, two agents acting the same will have to be judged differently as regards their moral worth: He who acts consistently over time will be judged as more morally praiseworthy than the other, he who has feelings of benevolence and care will be judged as more moral than he who acts solely out of justice, he who feels genuine concern for the well-being of others will be judged as morally better than he who has no inclination to help other than the respect for his duty. In the group of borderline cases mentioned above, i.e., when the morally right act loses its significance if the moral motivation behind it is one based on a sense of justice, duty or right rather than affection and care, one could argue that such actions are beyond the scope of morality but this distinction is far from obvious.

When the capacity for rational judgement and action is coupled with a mature affective life a further criterion for moral agency, namely autonomy also acquires a broader meaning. Some elements of the emotional maturity of an agent, i.e., the capability of correctly interpreting the emotions, thoughts or motivations of others, of having a thorough acquaintance with one's own inclinations, feelings and aims as well as a genuine and unaffected care and concern for the interests of others make it more probable to act rightly in a given circumstance. Kant's anti-consequentialist point that one can never predict the consequences of one's actions notwithstanding, it is obvious that one can be better or worse equipped to take the possible consequences into account and such worldly wisdom seems to be a practical part of autonomy.

For acts can turn out to be morally right by chance; an agent can happen to hit the mark without knowing or without even intending it, as a coincidence or when she acts under the guidance of somebody else. There is no little temptation to prefer the latter case: An act is comfortably heteronomous – or *unmündig* – if one is incapable of using one's understanding without the direction of others (Kant 1974, p. 9). It

would be unfair on Kant to claim that his understanding of autonomy disregards affective components; on the contrary, one of the main reasons that many people – including all females (ibid.) – remain under the guidance of others is their preference for comfort and their lack of courage. Nevertheless, the Aristotelian conception of the autonomy of a virtuous agent goes beyond and gives the character of the moral agent with its cognitive, affective and motivational aspects its rightful place. It is not enough to act rightly and for the right reasons, one must also have the right emotional attitude towards one's self, one's actions and just as importantly towards others.

The last but arguably most elementary criterion is one that both involves all those mentioned above and additionally is a fundamental characteristic of any postulation of agency in itself. Both conceptually and theoretically, it seems that the idea of agency requires some sort of unified self and it is this alleged relation between moral agency and self that sets a host of new questions. The first difficulty in an attempt to answer these is the need to gain some conceptual clarity; to understand what is meant by self, personhood and agency. This seems to be a necessary preliminary step, not only because of the intricate nature of the underlying subject-matter, but also because considerations regarding it have often dwelled round related termini, often used synonymously or interchangeably. Therefore, approaching the relation of self and agency first and foremost involves a conceptual clarification of the terms involved and a subsequent analysis of their relations.

A further set of questions opened up by such an enquiry into moral agency and self include the question whether a moral self exists, what its nature is and in what relation to the “self at large” it can best be understood. Apart from such a more “internal” survey however, is the trickier question of which types of agents should be

understood to have such a self. This train of thought takes us to quite concrete questions about collective action and collective responsibility and it is not clear how one can deny that collectives can be culpable – as collectives, not as a conglomerate of individuals – but it is just as unclear how one can postulate this. In this, as in the other problematic cases of moral agency we have encountered so far, there does seem to be differences in degree, which need to be taken into account. A collective of criminals, for instance, gathered round a specific common aim will probably not provide a court with its most difficult case as to individual responsibility and individual sentences. A nation or a cohort on the other hand will prove to be much more resistant to be regarded as a unit of self, with a collective aim, a collective purpose and collective responsibility. It seems that the relation of moral self and collectives requires a much broader understanding of identity than the one, which one tends to apply to individual human beings. On the other hand, it might turn out that such a broader and disjointed understanding of identity might be much more appropriate for individual human beings than the unified, single self that has proven to be popular in these parts of the world. If this is the case, then even the traditional postulations of a self, that consists of parts, such as the ancient one with a couple of parts of the soul, that of the Abrahamic religions with some parts infested with sin or modern ones such as the Freudian tripartite psyche seem inadequate. A wider understanding of the moral self, at the same time more intricate and broader than those mentioned above seems called for and might provide new insights into old problems of inconsistency and *akrasia*.

On the whole, it seems that a comprehensive understanding of moral agency requires most or all of the criteria above and both philosophical and sociological enquiries into individual agency focus on one or more. However, for some, their

being focused on the “individual” is their very problem. The discourse on agency and self, especially in the Anglo-Saxon tradition relies heavily on individualism, which often neglects many socio-psychological aspects of agency and sometimes even openly rejects them. One thus encounters queer sorts of creatures, particularly as the agent of Rational Choice Theories; some being covertly advertised to be endowed with a variation of a Kantian holy will, purely rational and quasi omnipotent, even if not morally impeccable. In its crudest versions, Rational Choice Theory postulates a straightforward relation between choice, agency and will: The will of the individual agent is bent on making the most rational choice in a given situation (Barnes 2000, p. 17).

Others are more wary of postulating a rational individual as the focus point of agency. One very obvious line of objection against such atomism comes from a tradition of sociology and political theory, in which the individual is regarded as essentially and necessarily embedded in her social environment and not “metaphysically independent of society” (Taylor 1985, p. 8). This amounts to more than claiming that the (moral) character of an individual is strongly influenced or even determined by her environment. The idea that “an individual is constituted by the language and culture which can only be maintained and renewed in the communities he is part of” (Taylor 1985, p. 8) has become as common as purely rationalistic theories of motivation and agency.⁸ Further criticism at the anthropological premise of these theories come from empirical studies on habitualised patterns of thought, opinion, behaviour and even aesthetic taste, which demonstrate high correlations between social origin, habitus and life-style (Bourdieu

⁸ This view is clearly distinct from subjectivism, which is often – and often mistakenly – attributed to so called post-structuralist lines of thought; “[...] the disengaged identity and the designative account of meaning it [subjectivism, B.S.] gravitates toward centres everything on the subject, and exalts a quite unreal model of self-clarity and control” (Taylor 1985, p. 11).

1982). An even more challenging point of view of the self claims that it is impossible to talk of a self in and by itself; self is necessarily and inevitably relational. Just as it is nonsense at worst or incomplete at best to talk of two-place predicates like 'near' or 'far' without specifying the relational point of view, it is just as misplaced to talk of a self existing on its own (Buber 1995).

Yet another line of objection meets with even more resistance than the first; probably owing to its being part of the Darwinian version of the blows to human narcissism. After having been forced to grudgingly accept that the Copernican view of the solar system seemed to accord better with empirical data and a short time before the popularity of Freudian psychoanalysis forced them to review their self-understanding as masters in their own house, Darwin struck at the flattering picture many Western thinkers had of themselves as made in God's image. Even if more modest thinkers might have been unsure about the alleged affinity to God already before Darwin, there was little doubt generally that human beings were essentially different from other animals. Animals, indeed, are still the untiringly reproduced other of the rationally and intentionally acting human being endowed with reflection and language. The history of philosophy is full of examples of the quest for the decisive property that distinguishes human beings from other animals, the most popular candidates naturally being rationality and language. Even today, it is still more the rule than the exception to encounter claims about this or that unique ability of humans; human beings are the only beings who can have second order desires (Harry Frankfurt, quoted in Taylor 1985, p. 16), they are "unquestionably unique in the shape and quality of [their] moral experience and behavior" (McGeer 2008, p. 117) only human beings experience sexuality, only human beings can experience

music etc. We will see that not even Buber is free from this when he claims that only human beings are capable of I-You relationships.

Harry Frankfurt points in a similar direction as Charles Taylor when he discusses what qualities are particular to personhood. They are not simply qualities that distinguish us from other creatures, “non-persons”, so to speak but rather “they are designed to capture those attributes which are the subject of our most humane concern with ourselves and the source of what we regard as most important and most problematical in our lives” (Frankfurt 1998, p. 11). What matters is not, therefore, what is specific to the human species. What matters according to Frankfurt is rather whether an individual is capable of forming second-order desires and he postulates that no other animals than human beings can form more than so-called desires of the first order. This differentiation between first order and second order desires aims to capture the difference between having desires, beliefs, motives and the capacity to make choices on the one side and the desire or the will to have some desires and motives or to lack others. What distinguishes human beings is accordingly the structure of their will, which is capable of “reflective self-evaluation” (Frankfurt 1998, p. 12).

I agree with Frankfurt when he emphasises that what is philosophically interesting about personhood does not have to be species-specific. It is, in fact, irrelevant which species possess the capacity to have second-order desires. It is a contingent fact that “No animal other than man, however, appears to have the capacity for reflective self-evaluation that is manifested in the formation of second-order desires” (Frankfurt 1998, p. 12). I shall therefore refrain from approaching the question considering self or personhood via the attempt at establishing what distinguishes human beings from other animals. More interesting is the link that both

Taylor and Frankfurt postulate between the capacity to evaluate, i.e., between making choices among one's own motivating drives. Selfhood in Taylor's terminology and Personhood in Frankfurt's terminology then is not something independent of the values one harbours and brings into actions. A person does not simply happen to have a strong evaluation, in Taylor's sense of certain courses of actions rather than a weak one. Rather it's her ability of evaluating herself reflectively, thereby postulating a divide between the self that is moved by first order desires and one that watches and directs that first order self.

The spectrum of these desires as Frankfurt understands them is fairly broad. It includes unconscious wishes to pursue a certain action, self-deception or conflicting desires:

To identify an agent's will is either to identify the desire (or desires) by which he is motivated in some action he performs or to identify the desire (or desires) by which he will or would be motivated when or if he acts. An agent's will, then, is identical with one or more of his first-order desires (Frankfurt 1998, p. 14).

But the will, as Frankfurt understands it, is that which is discernible by the action that she finally ends up doing. Thus this way to determine what the will is incorporates *akrasia*, self-deception and other puzzling phenomena between motivation and action. Conversely, it seems to fall into the danger of being trivial and uninformative. If we are to determine what the will is by retrospectively identifying it as that, which led to the agent to be moved or inclined and the action to be performed, the phenomenologically and action-theoretically questions are only answered by disappearing.

Frankfurt therefore discusses more interesting cases in which an individual wants to have a genuine desire without, however, wanting that desire to be effective, i.e., to be turned into action. A physician working with drug addicts might genuinely

want to be moved by the urge to take a drug and at the same time not want that urge to be satisfied (Frankfurt 1998, p. 15). Thus his second order desire leads him to want something the fulfilment of which he does not want. It's the wanting he is after and that is precisely what second-order desires are about. But "his desire to have a certain desire that he does not have may not be a desire that his will should be at all different than it is" (Frankfurt 1998, p. 15). There is another sense however and these cases are more common in that the agent wants to have a certain will and be brought to it to perform certain actions. These are the cases with which the question concerning *akrasia* deals and are the more straightforward examples of motivation linking to action, i.e., of his will as Frankfurt understands it.

It is these cases, i.e., not simply second-order desires, which are directed at having a desire but those that want a desire to be the will that are essential for personhood (Frankfurt 1998, p. 16). Such agents who, additional to first-order desires have second-order desires of the former kind but not of the latter, which Frankfurt calls *second order volitions*. Persons are such beings who are capable of having second order volitions while other beings who have first order desires and possibly also second order desires of the first kind are, in the terms that Frankfurt introduces, "wanton" (ibid.). Wantons are "all nonhuman animals that have desires and all very young children" (Frankfurt 1998, p. 16), possibly some adult human beings. Wantonness can come in degrees; one can be more or less wanton. Since wantons are characterised by not caring about their will,⁹ it appears that those second-order desires that do not involve self-reflective evaluation do not make a person out of an agent. It is not the lack of rationality or deliberation that makes an agent wanton; wantons are perfectly capable of reasoning about their wishes: "What distinguishes

⁹ "[...] it never occurs to him [the wanton, B.S.] to consider whether he wants the relations among his desires to result in his having the will he has" (Frankfurt 1998, p. 18).

the rational wanton from other rational agents is that he is not concerned with the desirability of his desires themselves” (Frankfurt 1998, p. 17). In this, Frankfurt’s position is comparable to that of Charles Taylor, who emphasises the importance of having strong evaluations for the existence of selfhood. The wanton is neutral about which of her desires should prevail and therefore I argue that the weak evaluator of Charles Taylor and the wanton are both (what I would prefer to term) amoral in the widest sense, i.e., in the sense of an evaluation that comprises both morality and the aesthetics of existence. Frankfurt emphasises the link of these second order volitions with selfhood with expressions like “identifying himself”, “making one of them more truly his own” or “withdrawing from the other” (Frankfurt 1998, p. 18).

Moral Patients

A common approach to answer the question of what makes up a moral patient is the reciprocal view of morality, which is also mirrored in contractual approaches to ethics and political philosophy. The idea that moral patients are extensionally identical with moral agents and are patients qua their qualities as agents has a long history, starting with the stoics and reaching out to Rawls via Kant:

In the same vein as Kant, but talking of rights rather than duties, contemporary authors have said that the right to life belongs only to moral agents, or to those who can possess the concept of a self, or to those who can reason, or to members of a natural kind whose mature members are normally rational. Similarly it has been said that we owe strict justice only to those who have a sense of justice” (Sorabji 1993, p. 129).

The idea here is that for someone to be harmed by the withholding of a good, they must have a conception of that good as such.

Determining moral patients only by their qualifications might involve the danger of overlooking other criteria of how a moral patient ought to be treated by a

moral agent: “In particular circumstances, we may be bound to particular animals, and the considerations which bind us are many, not just the one consideration of the position of animals in a series of concentric circles” (Sorabji 1993, p. 127). This is another example for the view that concentrating on the right moral action, i.e., in such a case, what sort of a treatment the moral patient should receive, neglects other morally important aspects, which deal with the relation between the moral patient and the moral agent and with the attitude of the moral agent more generally. The difference between a moral agent treating a moral patient in the right way but without the right reasons, the right motivation or the right emotional attitude, shortly, as a coincidence, an accident or because of coercion and another moral agent, who is in a particular moral relationship with the patient is a morally relevant difference.

A first look indicates that although the search for the criteria of moral patiency concentrates around a number of closely related termini, there are two distinct approaches. One of these two approaches is the traditional essentialist one: Entities are distinguished by certain essential properties, which make them what they are. What behaviour is expected from them – i.e., their position as moral agents – and what kind of treatment they can rightfully claim – i.e., their position as moral patients – depends on these essential properties. Historically, it was this approach that generally carried the day and various civil movements approached the reigning discourse deconstructively to show either that these properties were misattributed or that they were irrelevant. The latter approach focuses not so much on the essential or accidental properties of an entity as on whether or not it has interests. Considering the various liberation movements of the past couple of centuries from abolitionist campaigns to demands for universal suffrage or student rebellions, one is thoroughly

acquainted with this approach by now its underlying principle of equal consideration of interests is still widely contested.

Nevertheless, the non-essentialist, interest-oriented and egalitarian approach is just as bent on finding the relevant conditions for having interests, as the essentialist approach is on finding relevant essential properties. Of the former group, Jeremy Bentham famously insisted that the only possible criterion for having interests taken into account and thereby for moral patiency is sentience. His frequently quoted passage on the subject denotes the necessary quality as “the capacity to suffer” (Bentham 1996, p. 283) while his contemporary fellow-utilitarian Peter Singer speaks more broadly of “sentience”: “a convenient, if not strictly accurate, shorthand for the capacity to suffer or experience enjoyment or happiness” as the quality, which makes a being have interests (Singer 1986, p. 222). Charles Taylor avoids the question of which capacity underlies interests by referring to them directly under the term “significance”: “the fact that we are beings to whom things matter” (Taylor 1985, p. 2). Others rely on the concept of rights and argue that there are certain inalienable rights shared by all subjects of life, i.e., by all living beings, which have desires, a sense of self and an outlook into the future, which must be taken into account equally (Regan 1985), and if necessary enforced by law (Wise 2000). What unites these views is that the traditional view that this or that quality – most commonly rationality – is in itself necessary and sufficient to provide its possessor with a higher standing regarding the consideration of her interests gives way to a more egalitarian focus on having interests, regardless of any further qualities a being might have.

Which sort of beings have which sort of interests and which are to be given precedence over others in cases of conflict, are among the questions set by the non-

essentialist approach and the first difficulties with it are methodological ones. Some of the candidate criteria refer to empirical questions, some to conceptual, many to both. The question of which beings feel pain, for instance, belongs to the subject matter of a number of disciplines, e.g., philosophy, biology, psychology or neurology. In themselves, the questions what pain is or how it can be detected are not normative enquires but the subject of moral patiency appears when the quest for the criteria is united with a normative concern of how to treat beings endowed with the relative quality. The moral demand that underlies these principles seems to be the same throughout: Equality as a moral ideal; equal consideration of the interests of moral patients, independent of irrelevant factors such as race, sex, age or species on the one hand and – even more contentiously – the criterion of moral agency on the other. To have the actual or potential capacity for rational thinking or self-control may be a necessary criterion for being held accountable for one's deeds but it is difficult to see how it could be used to argue that the treatment of beings without these skills is not a moral question in itself. And yet, the idea that agency and responsibility are conceptually related – perhaps even interdependent – seems to be partly responsible for the long tradition of belief in reciprocity as a moral ideal.

One further reason for the persistency of this reciprocal view of morality might be the notion that responsibility is linked to the alleged tight relation between duties and rights (Simmel 1906). For some thinkers, the two terms are conceptually inseparable; one necessarily involves the other. If a person can be said to have rights, then someone is bound to have duties towards her. Both parties might be united in the same person; one can be arguably have duties and responsibilities towards oneself as the carrier of rights yet in more straightforward cases the two parties are separate. Faced with a moral patient who is attributed rights, the quest for the

'responsible caretaker' once more takes us into the heart of the question of moral agency and patiency. If for every right there must necessarily be some accountable entity responsible for providing for the right, then this can be a variety of types: Another person, a group of persons, an organisation or – commonly used as a justification for government – the state.

In the context of this line of thinking, one can only claim rights if one is ready to meet her responsibilities. A typical example of the outcome of such a postulation of a tight bond between rights and duties is citizenship. The underlying presumption demands that every participant of an ordered society is capable and willing to put at least as much into the community as she receives from it. There do seem to be some legitimate exceptions regarding age; potential citizens can rightfully demand services that will enable them to turn them into active citizens in time when they will be able to pay their dues. Similarly aged citizens, no longer capable of work, will be able to rightfully demand services for what they contributed to the commonwealth. Needless to say, such exceptions are the result of long and hard struggles in the establishment of rights that are granted citizens in a welfare state. However, historically, it appears that what is regarded as a full citizen is directly connected to what, in a given society, is regarded as an agent.¹⁰ It seems safe to claim that this bond is one manifestation of the reciprocal view of morality. Historically, consistency is not the greatest strength of this view. Women, for example, although they were not deemed capable of taking part in democratic governance, were nevertheless regarded as responsible and accountable for their deeds. So, while a woman found guilty by a jury of just and impartial fellow-citizens could easily have been sentenced to death, she was

10 Compulsory military service for able young men of a given society is a good example of which qualities of citizens qualify for first class political agency; in this example bodily intactness, age and sex. Obviously it is not a coincidence that these are among the most popular criteria used as justifications for discrimination.

nevertheless considered insufficiently knowledgeable and responsible to be given the vote.

After having struggled with sexism and racism more or less successfully, most liberal communities today have reached a point which is doubtless a much more improved one regarding individual rights than city-states of Classical Antiquity or Western democracies of the eighteenth and nineteenth century. One can hear Machiavellian tones even from David Hume, when it comes to the question of considering the interests of beings of inferior strengths. Animals, slaves and women, being creatures which share in the rationality of white adult men are nevertheless in no position to expect justice, since there can be no threat of retaliation or resistance from them; “we should be bound by the laws of humanity to give gentle usage to these creatures [hypothetical creatures, which are rational, but inferior in strength of body and mind, B.S.], but should not, properly speaking, lie under any restraint of justice with regard to them, nor could they possess any right or property, exclusive of such arbitrary lords” (Hume 1907, p. 23).

Historical inconsistencies apart, however straightforward this conception seems conceptually, there are intractable problems involved. The reciprocity demanded by such an understanding of rights and responsibilities regards the parties involved as more or less equal participants in a joint venture. This means not only that he who is provided with some common good by society is expected to contribute to the reproduction of the existing order but also that he who for one reason or another is incapable of providing these offices can be excluded from the number of eligible recipients of rights and services. In a social order which is based on a foundational myth like the social contract, moral agents and moral patients are

extensionally identical. Only somebody who can play his part as expected is in a position to have her rights respected.

Returning to our initial question regarding the boundaries of the moral community and the relations of its inhabitants, it seems that the traditionally popular views both leave a number of questions unanswered and additionally seem to be burdened with inconsistencies in themselves. Flattering as the Abrahamic view of man is, the sort of self, which is incorporated into divine law seems to expect too much of ordinary human beings and at the same time leaves out other potential candidates for moral agency. The similar anthropological view of man as being torn between a rationally willing self and an irrationally desiring one seems to be only a slight improvement and narrows the sphere for moral agency even further. With all the sociological and psychological challenges to the traditional understanding of self and identity, it seems that a comprehensive discussion of moral agency must first and foremost tackle the question of what kind of a self must be postulated to prevent the community from moral agents from becoming at the same time too wide and too narrow. Therefore, in the following chapter, I shall explicate some conceptions of self and some of the challenges posed by the study of pathologies of the self.

Compared with the difficulties that the study of moral agency involves, an enquiry into what qualifies a being for moral patiency seems relatively straightforward. Both the utilitarian approach of demanding equal consideration for interests of sentient beings as well as the rights-centred approach of demanding equal rights for “subjects of lives” (Regan 1985) seem more promising than the line of reciprocal contract morality. Working along the lines that these approaches offer might in time even accommodate deep ecologist considerations and expand the community of moral patients.

For, other than in the case of moral agency, the danger of misidentifying moral patients lies in keeping the community of eligible beings too narrow than too broad. There is certainly a risk in attributing some entities more moral agency than they are actually capable of, especially considering that due to law enforcement these might be claimed to be more responsible than they are and punished accordingly. The juridical system has in time accommodated such concerns and it is by now regarded as morally right to give a possible culprit the benefit of the doubt. Applied to the case of moral patiency, a similar principle necessitates us, when in doubt, to treat liminal beings with more consideration for their interests than less. The risk incurred in deciding whether or not the interests of a being are to be taken into account or if a being even is capable of having interests, is a risk taken in the name of the being in question and it is morally preferable to err on the side of caution.

CHAPTER 3

CONCEPTIONS AND SENSES OF SELF

This chapter will approach the question of existing conceptions of self. They will include examples of how the self was conceptualised in the history of philosophy as well as first person, phenomenological conceptions of self. The first starting point in this part is to differentiate various interrelated termini: “Personal Identity”, “Self”, “Soul”, “Mind”, “Consciousness”, “I” etc. At times these termini seem to be exchangeable; at times they refer to differing traditions of thought behind them (e.g., the term “personal identity” is generally used by contemporary Anglo-Saxon philosophers). I shall exclude a systematic survey of historical aspects of theorising (or experimenting) about the self and will only include cases, which are relevant to the question at hand. The fact that there are historically different strands of thought about self and habits of thinking about self and morality indicates that an important part of the self-conception of moral agents is culturally, socially and historically influenced.

Conceptual Disambiguation

The relations between these terms are problematic, as seems to be the relations between philosophers advocating the preferential use of one or the other. When for example Galen Strawson discusses the phenomenological basis of metaphysical postulates about the self, one point of critique he is faced with is terminological;

Kathleen Wilkes criticises his use of the word “self” and argues that what he “wants to do with the notion of a ‘self’ can be done equally well, and more economically, by the first-person pronoun” (Wilkes 1999, p. 25). Olson rejects the idea that there is a problem of the self, since there is not even the idea of the self to be problematic (Olson 1998). He takes the observation that no two people agree about what characterises the self to be an indicator that there is no actual subject matter of the self, the concept simply does not exist. What is being discussed under the term is so diverse that it would be better to do without the term altogether and concentrate on the different matters that are actually being discussed (Olson 1998, p. 645). All three traditional ways of conceptual approximation; definition, listing characteristic features and pointing at paradigm cases are blocked when it comes to the self, Olson argues (Olson 1998, p. 647).

It is true that conceptual ambiguity makes it difficult to be clear what the subject under discussion is precisely. Additionally it makes it easier for hidden premises to creep in. When Olson turns to the question of what a self would be, if there were such a thing, one of the first claims he makes is that “it is part of the meaning of the word ‘self’, in its typical philosophical uses, that a self be someone’s self. (That much, anyway, seems clear.)” (Olson 1998, p. 646). In fact, whether one is one’s self or whether one owns one’s self is among the cardinal questions concerning the self, as Olson himself considers later on (Olson 1998, p. 648).

The fact that there is no consensus on whether something exists or what its characteristics are, if it does exist is sometimes used as an argument to deny its existence altogether. Olson seems to argue in a similar vein as more familiar arguments from meta-ethics, which take the absence of agreement about whether moral facts exist to demonstrate that there are no moral facts. The argument from

disagreement does not work to demonstrate that what is being contested doesn't exist however. As in the case of other contested entities – God, aliens, or ethics for example – whether there is little, a lot or no agreement about their existence serves neither to prove nor to disprove anything. Olson's argument does strengthen the view however that the conceptual difficulties around the concept of self and related terms must be tackled.

Olson emphasises his point by comparing the term 'race' to the term 'self'. Although the former has been criticised for being vague, there are nevertheless key features that makes it possible to use it in sensible discourse, he argues. According to his position 'race', as problematic a term as it may be, is able to provide characteristic features (in the sense that prototypical examples have certain characteristics that distinguish them from other prototypical examples), consistency of application (practically everybody agrees on these features) and paradigm cases (prototypical examples serve as paradigm cases; certain individuals can be pointed out to fit the bill) (Olson 1998, p. 652). Thus, the case with the term 'self' is not comparable to other problematic terms although there is, in both cases, "a social problem about the word 'self'. There is no philosophical problem about selves." (Olson 1998, p. 652).

If the term "self" is such a conceptual non-starter, why prefer it to such other terms like "I", "Ego", "Subject" or "Person"? One reason is the terminological landscape that came to be historically. These terms, although they are so closely related that some doubt the legitimacy of their usage (Olson 1998) have each of them a specific connotation. "Ego" for example has been used by Freud and Sartre. "Person" is related to two distinct discourses, the analytical discussion on personal identity and the legal-ethical discussion of moral considerability. The "Soul" has

fallen into disfavour,¹¹ mainly because of its religious connotations, although there continue to be theorists who persist in claiming that the “Soul” is a useful concept.¹² The term “Subject” is distinctly continental. One advantage of the term “Self” is that it seems a little less historically tainted than other candidates.

While Olson argues that there is so little agreement about what the self is supposed to be as to make the philosophical problem disappear along with the concept, Flanagan argues that psychologically, what the self needs in order to become the foundation for (moral) agency is “uncontroversial”:

We can summarize the uncontroversial links among the psychological capacities subserving a rich identity, certain other valued capacities, and particular kinds of social conditions as follows: (1) self-representation is necessary for self-esteem, self-respect, and self-knowledge; (2) a firm and invigorating sense of identity is necessary for effective agency; and (3) all these things – self-knowledge, self-respect, an invigorating sense of identity, and effective agency – require for their development early social relations of a certain qualitative kind. (Flanagan 1991, 141).

Especially relevant for our purposes is point 2:

Without the invigorating sense of self, there is no person and thus no coherent cognitive and motivational core from which the individual can generate purposes or in which he can find the energy required to sustain them, were he able to find any in the first place (Flanagan 1991, 140).

Flanagan’s position is representative for the idea that a stable sense of self is a necessary condition for (moral) agency and for the idea of a core as the gravitational point of the forces of the person. Neurologically at least, there seems to be no place where “all comes together”, but is it otherwise true that effective agency would be

¹¹ Some argue that the soul along with its immortality and even near-divinity crops up under different guises; William James for example criticises the Kantian ego as a “cheap and nasty edition of the soul” (James 345)

¹² Ian Hacking for example rejects the idea of the soul as transcendental or immortal but considers it to be a good term to invoke “character, reflective choice, self-understanding, values that include honesty to others and to oneself, and several types of freedom and responsibility. Love, passion, envy, tedium, regret, and quiet contentment are the stuff of the soul [...] I do not think of the soul as unitary, as an essence, as one single thing, or even as a thing at all. It does not denote an unchanging core of personal identity” (Hacking 1995, 6)

impossible if there were no core, neurological or psychological? Many events, which one generally considers to be actions, take place over a long period of time, e.g., writing a dissertation. It does seem indubitable that there has to be constancy in a number of cognitive and motivational factors in order to accomplish such examples of agency. It does not seem as evident that the links between these at a given moment in time or those over a period of time justify talk of self as an entity with necessary synchronic or diachronic identity. Nevertheless, the idea that cognitive and motivational constancy or some sort of 'relatedness' is necessary for agency can't simply be shrugged off. The difficulty is in explaining how such inner coherency can subsist without reverting to a substantial account of the self.

Kinds of Conceptions of Self

Perhaps the best known critic of such a substantial account of the self was David Hume. Self, as he understood it was a bundle of psychological and/or biological states without an owner with rather the metaphor of a commonwealth fitting the bill than an individual.¹³ He notes that “There are some philosophers who imagine we are every moment intimately conscious of what we call our SELF” (Hume 2009, p. 393) but goes on to write:

But setting aside some metaphysicians of this kind, I may venture to affirm of the rest of mankind, that they are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are

¹³ Strawson refers to a passage in the introduction of *A Treatise of Human Nature*, in which Hume writes of the mind that its essence is unknown to us (Hume 2009, p. 12) and argues that such an agnostic view is Hume's basic ontological position, rather than a positive postulation of the bundle theory. The passage in question is concerned with a comparison of methods and the success of natural vs. so-called moral sciences however and quoted fully does not support Strawson's interpretation: “For to me it seems evident, that the essence of the mind being equally unknown to us with that of external bodies, it must be equally impossible to form any notion of its powers and qualities otherwise than from careful and exact experiments, and the observation of those particular effects, which result from its different circumstances and situations” (Hume 2009, p.12). Hume's point here is methodological rather than ontological.

in a perpetual flux and movement. Our eyes cannot turn in their sockets without varying our perceptions. Our thought is still more variable than our sight; and all our other senses and faculties contribute to this change; nor is there any single power of the soul, which remains unalterably the same, perhaps for one moment (Hume 2009, p.395).

What Hume rejects however is already a particular conception of self. Sorabji

questions why Hume took the self, which he demonstrates not to exist, to be

“something internal and introspectible” as well as something already existing rather than “something to whose creation, as Locke allowed, you might yourself contribute” (Sorabji 2006, p.18).

Similarly and in a characteristically bitter tone Nietzsche rejects the idea of the figure of a self behind the agency when he dismisses something akin to the self as a neutral foundation of the strong individual: “But no such substratum exists; there is no ‘being’ behind doing, acting, becoming; ‘the doer’ is merely a fiction imposed on the doing – the doing itself is everything. [...] our entire science is still subject to the seduction of language and has not shaken itself free of the monstrous changelings, the ‘subjects’, foisted upon it” (Nietzsche 1996, p. 30). This is the position that Derek Parfit calls “Reductionist”; it’s the denial that “the subject of experiences is *a separately existing entity*, distinct from a brain and body, and a series of physical and mental events” (Parfit 1986, p. 222).

In *Reasons and Persons*, Parfit himself rejects the Cartesian Ego, i.e., the idea that there is a thinker behind thoughts and a doer behind deeds, which is a substance of its own. He argues that it is possible to give an “*impersonal* description” (Parfit 1986, p. 225); i.e., a full description of all experiences along with the connections between them without the need arising for postulating a separate entities as the subject of these. Sorabji reports Parfit to have moved away from the idea that “reality could be fully described in impersonal terms: that is, without the claim that people exist” (quoted in Sorabji 2006, p. 266). Sorabji argues against Parfit that there is a

person as the owner of experiences and that such a person is a necessary condition for agency and therefore for ethics but avoids Cartesian dualism by considering that owner to be an embodied person (Sorabji 2006, p. 267).

Sorabji conceives of the person as the owner of psychological states and actions rather than a stream of consciousness (Sorabji 2006, p. 265). He follows Marya Schechtman in arguing that “duration is required for talking, listening, walking, acting, having beliefs, desires, goals, intentions, thinking, being inconsistent, and vacillating” (Sorabji 2006, p. 270). As a further argument he adds Schechtman’s point that “one experience is pregnant with others” (Sorabji 2006, p. 271). For Sorabji this sort of pregnant interconnection “often involves” that there should be the same owner but as “often involves” is not “necessarily requires” and the importance of ownership remains unclear. Sorabji also postulates in passing the reverse of his point that interconnectedness involves ownership; i.e., he argues that that ownerless experiencing would undermine interconnectedness and that this is symptomatic of a pathology, i.e., Alzheimer’s: “This is not the only example in which a philosophical theory describes as normal what is actually found only in sickness” (Sorabji 2006, p. 271).

Another way of conceiving of the self is to regard it as a narrative. Daniel Dennett famously argues that it is a category mistake to try to understand the self as a thing in itself with discernible properties. He suggests that the self is an abstraction in the way that the centre of gravity is an abstraction; it is a theoretical tool to work with and as such a useful fiction. The self is a more complicated fiction than the centre of gravity not only because it involves narration but also because it involves a special type of narration, which favours a positive and unified account of oneself (Dennett 1992). Dennett himself favours a modular model of mind, in which

functional units are partly autonomous and partly inaccessible to each other and argues that phenomena like multiple personality disorders or split-brain patients provide indicators for this.

Similarly Pierre Bourdieu speaks of a biographical illusion, which is created both by the subject and society and which is profitable in creating a semblance of coherence, necessity and a seemingly intelligible line of cause and effect (Bourdieu 1985, p. 76). Bourdieu does not only claim that persons prefer such an intelligible story-line as the story of their lives and that there are social mechanisms, which support the experience and narration of life as a unity and totality (Bourdieu 1985, p. 77). Among such mechanisms for summing up and unifying the I is the personal and legal name along with all the rituals connected with it, the signature, with which properties of an individual can be transferred from one social field to another. Bourdieu argues that it is elementary to consider the matrix of the social fields through which the subject moves and the trajectory of which is constructed as her “biography”. To consider a subject without this social embeddedness is as “absurd as the attempt, to describe a trip with the underground, without taking into account the structure of the net” (Bourdieu 1985, p. 82)

The dangers of fabricating the story of one’s self and one’s life have been dwelt on quite often. The general idea is that the tendency is towards inventing a rosier picture of oneself, of distorting unflattering facts in one’s past, of belittling the importance of the qualities one regards as unfortunate and emphasising the importance of those that are accorded a high value. This sort of self-enhancement via narration serves to present oneself as a “morally decent, possibly as morally good, even virtuous” (Flanagan 2009, p. 65). Nietzsche is famous among those pessimistic about the tendencies of human beings of distorting facts to fit their view of

themselves. One historically/sociologically relevant reason why narratives have an important role in the assessment of others is that in the non-tribal way of life, narratives have perforce taken the place of direct observation of behaviour. Narratives as part of self-disclosure have gained weight in ascribing traits to others. Considering the idea that self-deception is the best way to successfully deceive others, it is also to be presumed that there is an evolutionary force to deceive oneself about the moral decency: “Philosophy, along with other critical disciplines, can help us examine what questionable factual, moral, or metaphysical assumptions our narratives make, embed, and enact” (Flanagan 2009, p. 75).

A further point about narratives, which ought to be kept in mind, is that they involve a number of assumptions about morally relevant philosophical questions. These can often be quite implicit; an unreflected, learned usage of assumptions common in a given discourse. Their being implicit is problematic in increasing the power of such assumptions over the arguments employed by the agent in establishing moral judgements or habitualising patterns of moral actions and reactions. This is one area where it is important that the metaphysical or meta-ethical assumptions made about free will, the distinction between moral and non-moral realms, responsibility or agency should be made explicit so that they can be turned into areas of reflection. This is especially important in moral education; moral education seen in this way must involve making implicit metaphysical assumptions explicit in what Flanagan calls “moral education as metaphysical critique requirement” (Flanagan 2009, p. 67): “Narrative” in this context is understood much more widely than the biographical narrative of an individual agent; it involves the story told by the agent to herself and others about the metaphysical structure of the world, of action and of

morality.¹⁴ Narratives are thereby never normatively neutral; they are “richly normative and give guidance and direction on how things will go from here, and on what is the likely trajectory, both empirically and normatively of this life or these lives” (Flanagan 2009, p. 69). Flanagan (Flanagan 2009) speaks of “master-narratives” (p. 54), which are commonly accepted in a given culture and which influence individually made up narratives.¹⁵ The urge to understand the self and more importantly one’s own self as coherent and unified would accordingly be deeply intertwined with views on morality and a general world-view.

Character and No-Character Theories

One aspect of the question concerning the unity of the self is the correspondence between attitude and behaviour and the overall stability and predictability of both. The philosophical question whether there is a stable, unified self; an agent who produces behaviour that is consistent over time and thereby predictable, is mirrored in psychology as the question whether behaviour is better explicable and predictable by invoking character, traits and dispositions or by situational factors. A contested

14 Types of narratives that are common in a given discourse thus have a great impact on the moral judgements and choices made by the individuals who buy into these: “A major function of master-narrative structures is to situate persons and lives in a moral space by depicting types of lives that are deemed decent, good, noble, virtuous, and the like. The patterns of familiar narratives allow us to quickly classify whether individuals are good or not, trustworthy or not, and what sort of karmic outcomes are likely to accrue in their vicinity”. (Flanagan 2009, p. 68): “Master-narrative” does sound a little as if there is a “master-plan” behind the reigning patterns of perception, taxonomy and evaluation in a given society. The postulation of master-narratives however is obviously quite neutral about the historical questions of how such master-narratives come to be prevalent. Theories of how such reigning patterns are established include theories of consent (Habermas), power (Foucault) or complex inter-dependencies between field and habitus (Bourdieu).

15 “When I speak about myself (or you), especially if I tell part of my story (or yours), I stand on the shoulders of ancestral storytellers who have supplied what are now – but once were contested – commonsense categories and familiar plot lines in service of the interpretation of persons and their lives. These ancestral storytellers were themselves dependent on communities of predecessors who invented and/or stabilized the language we speak, parsed the universe, and introduced work linkages, word spans, that attempt to capture what we now think of as our kind of beings-in-time doing what our kind of beings-in-time do in time.” (Flanagan 2009, p. 65).

point is whether traits are concepts, which can be used as predictors for behaviour or whether the language of character and traits is simply a manner of speaking, a manifestation of “our needs to encourage and discourage, praise and blame, reward and punish, ourselves and others“ (Sabini & Silver 1982, p. 143). Put differently, this is the question whether talk of 'character' (as a synonym for 'self' or understood as part of it) had better be understood as an explanatory and predictive concept of human judgement and behaviour or as only of that manifestation of human behaviour that tends to attribute stability, reliability and constancy to human judgement and behaviour. And if so – how stable a trait is that tendency to attribute character?

As in the case of self, there is no consensus on what character means. Is character the elusive ‘inner’ of a singular human being as opposed to the outer? Are traits quasi-perceptual mechanisms like those of sight or touch, which manifest themselves as stable patterns of reaction towards “outer” stimuli? Flanagan considers whether they are “tendencies to express reliably certain patterns of perception, feeling, thinking, and behaviour, similar perhaps to my know-how for bike riding, which is not in me as an area of my brain is in me, but is a disposition in me that is activated by bikes; and which is not possessed by my friends who don’t know how to ride bikes [...]” (Flanagan 2009, p. 53). Another possibility is that they are conceptual constructs to deal with recurrent phenomena in the relations between self and the rest, without ontological reality, in the way that sunsets are conceptual constructs without ontological reality (ibid.). A further possibility is that character traits as such do not exist at all; a view which has been prominently advocated by John Doris (Doris 2002). Flanagan criticises this “ontologically mischievous” view, “that character traits are like phlogiston or unicorns, and thus that moral theories that depend on the positing of traits – virtue theories first and foremost (in fact, the

criticism, if it were apt, would apply to all moral theories West and East) – are non-starters”¹⁶ (Flanagan 2009, p. 55). This is another example of the view that a self with some amount of unity and stability is seen as necessary for moral agency.

Flanagan’s psychological realism is more successful in explaining the stability and recurrence of patterns in judgement and behaviour. Indeed, it would be very difficult to explain how behaviour in the widest sense could stay independent of the two big determinants of nature and nurture. Even if we suspend judgement on whether nature and nurture exhaustively explain and predict (moral) behaviour, it would need a queer theory to explain how judgement and behaviour come to be localised in a given moral agent without these factors. The situationist challenge that Doris poses is not necessarily incompatible with the view that these determinants are effective in a given – situational – choice of action. Whether certain actions are “coded” in a certain, culturally, psychologically and sociologically influenced way that does not do justice to their specific character, does not mean that there is nothing to habitualised patterns of perception and reaction but might just be lazy taxonomy.

Jonathan Haidt proposes a model, in which both aspects are integrated. He argues that in the answer they give to the question of where moral beliefs come from, empiricists are mistaken in conceiving the human being as a *tabula rasa*, who acquires them via experience (Haidt 2008, p. 182 see also Pinker 2002). He also rejects rationalist accounts, which focus on reasoning and moral sense theories, which focus on the emotional aspect. His own model takes into account all of these

¹⁶ Flanagan also questions the use of the low correlation between traits and behaviour as supporting a situationist interpretation. Low figures of correlation between traits and behaviour tend to be interpreted as indicators either that traits as such do not exist or that they have little effect on behaviour. Flanagan however argues that the numbers have been misinterpreted. The sort of correlations that have been found in the past (around .30 or .40) seem to indicate that traits and behaviour are unrelated. If the chance of predicting the choice of action of a given subject S has an accuracy of 50%, then Flanagan argues that a correlation coefficient of .40 will raise this to 70%; which is quite a high figure (Flanagan 2009, 63). Both the disposition to act in a certain manner as well as the features of the situation contribute to the behaviour that the agent ends up producing.

factors but emphasises social intuitions. According to the Social Intuitionist Model of moral Judgement (SIM) the mind is “inescapably affective” (Haidt 2008, p. 187) and judgemental. The majority of judgements are captured by the psychological processes of intuitive judgement, post-hoc reasoning, reasoned persuasion and social persuasion. What is sometimes considered as the main path of moral judgement, i.e., reasoned judgement and private reflection are also given their place in the model but they are considered to have lesser roles (ibid.). In fact, they can sometimes be detrimental to good moral agency because they are often not the reason behind the judgement but rather the rationalisation after it: “In fact, this human tendency to search only for reasons and evidence on one side of the question is so strong and consistent in the research literature that it might be considered the chief obstacle to good thinking” (Haidt 2008, p. 190). Such reasoning can not only serve to justify the deed after it’s done, it can also shift the borders of the moral and the amoral: „We can persuade ourselves that, really, there is no ethical dimension at all to the situation in which we find ourselves“ (Fine 2005, p. 75).

These different possibilities of understanding character traits are mirrored in varying degrees of espousal or scepticism towards virtues and vices. Flanagan is an example for espousal; he argues that “Reference to virtues and vices, and to the aim of trying to equip agents with a good character comprised of virtues is psychologically, sociologically, and politically wise, as well as ontologically respectable” (Flanagan 2009, p. 56). According to this view, there is no reason to posit ontological arguments against character traits because traits are not things in the ontologically conservative sense, they are dispositions; patterns of correlations between certain situations on the one side and judgement and behaviour on the other.

As such, they are not susceptible to ontological arguments against them while providing functional working tools to understand and predict behaviour.

A suggestion of what an ‘ontologically respectable’ way of speaking about morally relevant traits (i.e., virtues and vices) is provided by Flanagan: “A virtue is a disposition {to perceive &v to feel &v to think &v to judge &v to act} in a way that is appropriate to the situation” (Flanagan 2009, p. 56). There is also room for moral theories that put more weight to one or the other of the five aspects put into the calculation in the definition above.¹⁷ One might argue that every moral theory has assumptions about how these components contribute to moral judgement and moral behaviour, i.e., a theory on moral psychology and additionally normative views on how the ideal configuration of these components ought to be.¹⁸ I agree with Flanagan who argues that “The upshot is that debates about the relative causal efficacy of traits versus situations is a discussion about the relative causal power of two kinds of causes, where both exist. There are traits, and there are situations. They interact. End of story” (Flanagan 2009, p. 64).

17 “Socrates and the Stoics did not think “feeling” was desirable in the activation of the virtues, whereas Plato and Aristotle think it is essential. Confucious and Mencius think we just need to grow the good seeds that are already inside us in order to become virtuous, whereas Mozi, who comes between the two, is said to think the mind is a moral tabula rasa, and thus that virtues like compassion and honesty will need to be built from scratch in the way my ability to play a musical instrument is (but see Flanagan, 2008). Hindus, Buddhists, and Jains all think that there are poisonous dispositions in our natures that require elimination in order for positive dispositions, the virtues, to take hold. Iris Murdoch, Simone Weil, Lawrence Blum (1994) emphasize acute, particularistic, perceptual sensitivity more than most ancients and, in part, because of the more complex requirements of modern social worlds. The virtues of the Buddhist bodhisattva or the Christian ascetic don’t require much in the action department, but Confucian and Deweyan virtues do. And so on.” (Flanagan 2009, p. 59).

18 Basing such descriptive and normative moral theories on dispositions might lead to “metaphysical anxiety” (Flanagan 2009, p. 61) in some, since the ontological status of dispositions might seem queer. However such anxiety is only justified in a fairly restricted ontological view, which concedes ontological status only to things and excludes processes and events. Dispositions are not primary qualities but predictions that given the ‘locus’ of the disposition and certain conditions, a certain series of events will follow. Thus “the essence of a virtue is to be a disposition designed to be situationally sensitive” (Flanagan 2009, p. 62).

Sense of Self

Ideas like those exemplified above are concerned with the question of what the self is; i.e., they approach the self from a third person point of view. The self along with its related aspects like consciousness, perception or agency is a unique object of study as having both a first person and a third person access to it. A comprehensive study of the self must therefore include the phenomenology of the self. The need to study the self both from the first and the third person views is not only grounded in the fact that there are these two types of access to it but also because the very capability of having an idea of what a self is will have an effect on the sense of self: “The topic of self-awareness is not entirely distinct from the topic of self-hood, because the identity of a person (is) seen (...) to depend partly on one's conception of one's self” (Sorabji 2006, p. 9). The sense of self is not independently given but influenced by the third person view of self – which in its turn can be nourished from as diverse elements as personal experience, education or reigning conceptions of self in a particular culture.

Tackling the question whether there is such a thing as the “self”, Strawson argues that one needs to investigate two or three preliminary questions about it (Strawson 1999, p. 2). The local phenomenological question concerns the nature of the particularly human sense of self, which in turn leads us to the general phenomenological question of what the minimal conditions are to possess a sense of self. Via the final, general question of the conditions of a sense of self, we may find ourselves in a better position to answer the question, whether there is such a thing as self. This approach seems unusual at first in that it combines the distinct questions concerning the sense of self with the one concerning the existence of the self.

The first question regarding the human sense of self presupposes that there exists a universal, shared sense of self regardless of culture (ibid..). Its nature is mental; “a single mental thing that is a conscious subject of experience, that has a certain character or personality, and that is in some sense distinct from all its particular experiences, thoughts, and so on, and indeed from all other things” (Strawson 1999, p. 3). This is a distinctly phenomenological, first person understanding of the self as an owner (Sorabji 2006) or carrier of mental states, not simply a collection or row of mental states. A sense of self acquired during childhood with the gradual realisation that one's thoughts are private and not directly observable by anyone else is, as Strawson argues, something very ordinary (ibid.).

What is the self experienced as? Strawson lists eight qualities, which seem to be primary in what most people explicitly think of or implicitly assume of their selves. An important exception to this allegedly universal set of qualities of a phenomenological sense of self are certain pathologies that seem to hinder the experience of one or more of these qualities, causing a sense of estrangement and alienation of oneself. Strawson's list is useful for analysing in more detail how these singular assumptions about what makes up our sense of self influence our sense of agency and thus our sense of ourselves as moral agents.

The mental self, according to Strawson, is

conceived or experienced as:

- (1) a thing, in some robust sense
- (2) a mental thing, in some sense
- (3, 4) a single thing that is single both synchronically considered and diachronically considered
- (5) ontically distinct from all other things
- (6) a subject of experience, a conscious feeler and thinker
- (7) an agent
- (8) a thing that has a certain character or personality (Strawson 1999, p. 3).

This description combines most of the traditional postulates concerning the Ego, the Self or the Person. It differs from them in that it makes a claim on what is popularly believed to be the self; i.e., the view that individuals have of their selves. It might be speculated that many if not most of the philosophical views on the nature of the self were nourished by what such phenomenological experiences and/or by what was the popular view. A more intricate postulation on the mutual influence appears when we consider the effect on both the popular view on the philosophers mind as well as the direct, immediate introspection.¹⁹

Some of the components in the list, such as the mental nature of the self are contested, while others open up more basic ontological questions on the nature of agency or 'thing'. For our purposes it is sufficient to be able to claim that these factors are indeed the most common, if not universal among self-conceptions. Which of the elements in the list above are considered as essential, might in fact be used to distinguish various strands of claims about the nature of the self. Strawson's own position tends to regard numbers 4 and 8 as not necessary to a sense of the mental self; i.e., a human sense of self might exist even if does not regard itself as existing diachronically and - relatedly - not as having something akin to a character (Strawson 1999, p. 4). One of the least contested is the minimal view of the self as the phenomenal consciousness subject; the core self, which is nothing but the subject of experience or, as sceptics will have it, nothing but the experience itself: "Most people have at some time, and however temporarily, experienced themselves as a kind of

¹⁹ In his latest book (Strawson 2009), Strawson coins the term 'sesmet' (subject-as-single-mental-thing) to denote the singular and mental being of selves (Nagel 2009, p. 33). "Philosophers have not been very successful in devising credible accounts of the identity of the self over time, and Strawson's arguments help us to see why this is so. It seems to require that a single mental subject should be capable of existing without any consciousness and through vast changes of experiential content, but it is not clear that the mere physical existence of the brain is sufficient for this, and an immaterial substance may be no better" (Nagel 2009, p. 34).

bare locus of consciousness - not just as detached, but as void of personality, stripped of particularity of character, a mere (cognitive) point of view” (Strawson 1999, p. 13). The 'cognitive' had better be replaced by 'perceptual', for among the best examples of such moments are artistic experiences, as producer or consumer.

Strawson claims that "the self isn't thought of as merely a state or property of something else, or as an event, or process, or series of events" (Strawson 1999, p. 8). Accordingly, the self must be regarded as a thing simply because there is no other ontological option left. At first sight it might seem as if it does not matter very much, whether the self had best be understood as a thing or rather as a process. The ontological problem here is the familiar one related to all entities that have an event-like character and for my purpose of investigating the relation between the sense and theory of self and the moral stance, the decisive point seems to be whether one thinks of oneself as a single and stable entity, regardless of whether that entity is more akin to a thing or to an event. Practically however, and possibly also conceptually, this distinction is significant. If the self is understood as something like a soul, then the implications for agency, responsibility and accountability are different than if it is to be understood as something like an ever-changing stream of mental events rather than a single thing.

While such singularity is sometimes considered to be a condition for agency, Kant conversely takes the singularity of the 'the soul' or 'the thinking I' to be grounded in singular action (Kant 1877, A 351). Christine Korsgaard follows him in this view when she claims that "it is essential to the concept of an action that it is attributable to the person as a whole, as a unit, not to some force that is working in her or on her" (Korsgaard 2009, xii). This unity of action is also given in the case of thoughts, as Kant tries to demonstrate with a *reductio ad absurdum*; if the

conglomerate would think, then every part would contain part of the thought and only all parts together would contain the whole thought. This however, Kant argues, is a contradiction and therefore a thought can exist only in a singular entity²⁰ (Kant 1877, A352). Kant further argues that although it might be possible to think of a thought as divided among many subjects, the subjective *I* itself could not be divided (Kant 1877, A354). Strawson exemplifies the argument in a more vivid manner; supposing that one were to learn that one's sense of self was the product not of one's brain but of the co-operation of a number of brains. The experience of singular subjectivity would not be affected by this (Strawson 1999, p. 10). The point made here is that the experienced singularity of the self is that of a mental singularity, not of a bodily singularity: "It is true that ordinary human experience of oneself as mentally single is deeply shaped by experience of having a single body, but it hardly follows that any possible experience of oneself as mentally single depends essentially on such experience" (Strawson 1999, p. 10). Strawson makes a similar case for the independence of the experience of a unified self over time by arguing that it would not matter for this sense of self to discover that it was produced by a multiplicity of physical entities and events (*ibid.*). All these musings are in fact about the relations between conscious experience and underlying ontological entities. Conversely, a case in which non-mental unity was given while singular, diachronic mental events lacked a sense of singularity with each other would indicate that there was a lack of singularity, even if all this took place in a single brain. According to this line of thought the sense of a singular self is that of a singular *mental* self.

20 Kant's argument is a forerunner of the Chinese Brain thought experiment: "Denn, weil die Vorstellungen, die unter verschiedenen *Wesen* verteilt sind, (z.B. die einzelnen Wörter eines Verses) niemals einen ganzen Gedanken (einen Vers) ausmachen: so kann der Gedanke nicht einem Zusammengesetzten, als einem solchen, inhärieren" (Kant 1877, A 352).

Examining the claim that there are cases of people who experience themselves not as singular selves but as multiple ones, Strawson argues that there is a trivial sense in which an experience of multiple selves is not possible. It is doubtlessly the case that many people report experiencing themselves as fragmented, torn, split or multiple. This sense of self as distinctly not singular is not only familiar from literature on mental pathologies but rather a basic aspect of the human condition, often subject of literary or artistic interest. Strawson concedes that there might be experiences where “[o]ne may be under stress and subject to rapidly changing moods. One may feel oneself pulled in different directions by opposed desires. Human thought-processes can become extraordinarily rapid and tumultuous” (Strawson 1999, p. 12). But he rejects the interpretation of such phenomena as indicating that there might be multiple selves since “the experience that there are many selves present is necessarily experience from some single point of view. Even if a single brain is the site of many experiences that there are many selves present, each such experience is necessarily experience from a single point of view” (Strawson 1999, p. 13). Accordingly, the experience of being torn between conflicting desires or ‘wills’ can only be experienced as conflicting if there is a single self experiencing them as one’s own (ibid.). The entity in question had better be understood as Buridan’s ass between two haystacks, rather than as two separate asses.

This is not a convincing argument however. What makes the various experiences experienced as comparable and (in the case of a feeling of being split among them) as conflicting is that one has access to them all. In this case the access in all instances is first hand, from the first person view. Taking this as a reason to argue for a singular point of view, i.e., a self experienced as singular might be

begging the question because although there is a sense of sharedness or commonness between possible multiple selves, i.e., the fact that there is a first person access to all of them, this does not mean that they cannot be understood as separate. Strawson argues “Nor can one experience conflict of desire unless one experiences both desires as one’s own” (Strawson 1999, p. 13), i.e., if and only if one experiences all desires involved as one’s own, can one experience conflict of desire. But there can be experience of conflict among the desires of distinct entities, the desires of which are both experienced from the first person view. To postulate that therefore there must be one and only one entity that has access to this perspective begs the question. There might be none at all, as Hume would argue. Or there might be many, having first person access to each other, so to speak.

Strawson criticises James's metaphor of 'the stream of consciousness', arguing that “[h]uman thought has very little natural phenomenological continuity or experiential flow [...]. It [our thought, B.S.] is always shooting off, fuzzing, shorting out, spurting and stalling” (Strawson 1999, p. 17). It is expedient to distinguish here between “consciousness”, “(trains of) thought” and “sense of self”. One obvious distinction is that a sense of self as a consciously held experience of oneself as such is not identical with either consciousness or thought. The relationship between these three concepts is problematical. Can there be thought without consciousness or consciousness without thought? Partly, this is a question of definition but in a more interesting sense it is more. Strawson’s criticism of the Jamesian metaphor of “flow” and “stream” is based on his observation that his own states of consciousness are ruptured moments of a “starting” of consciousness with moments of complete unconsciousness (Strawson 1999, p. 18). He argues that experiencing consciousness as a flow is partly the effect of having learned to think of it as a stream. The

semblance of continuity is not grounded in a phenomenological continuity of consciousness or conscious experience but in the continuity of content; i.e., the coherent storyline built up with the help of various cognitive mechanisms like short-term memory, confabulation etc. Strawson (1999) argues that this sort of constancy in content is confused with constancy in the flow of the mental self, while the latter is, in fact, sporadic, episodic and saccadic (Strawson 1999, p. 19), thus the “belief in the flow of consciousness may be one interesting and suspect source of support for a sense of long-term continuity” (ibid.).

Strawson’s own view of the self is comparable to the Buddhist conception of a self, which does not persist over time, i.e., which lacks the conditioning of diachronic unity. Strawson calls his own view the “Pearl” view; the metaphor denotes a number of mental selves as lined on a string,²¹ where consciousness is the main binding criterion and regards as the basic unit of mental self each epoch of consciousness. Strawson (1999) writes:

[T]he mental self – a mental self – exists at any given moment of consciousness or during any uninterrupted or hiatus-free period of consciousness. But it exists only for some short period of time. But it is none the less real, as real as any rabbit or Z-particle. And it is as much a thing or object as any G-type star or grain of salt. And it is as much a physical thing as any blood vessel or jack hammer or cow (p. 21).

The Self and the Brain

One way of searching for the elusive self as the basis of moral agency is to take a materialistic stance and look for it in the body. Not that any part of the body provides a straightforward locus for the self; even from the relatively simple point of view as a

²¹ Strawson has later said that he regrets having used the metaphor of a pearl since he rejects the idea of there being an underlying “string”, whether understood as a stream or a narrative (personal communication, 23 November 2010).

physical entity, a human being is more a hybrid than simply a human being – the ratio of human cells and microbes that make up this symbiotic physical system is 1:10, i.e., only 10% of the cells in a human body are human. The rest are alien, predominantly bacterial cells, which are vital for the whole organism (Anderson 2008). Different disciplines like philosophy, psychology and biology, which deal with various aspects of consciousness and behaviour, have been converging on the view that the brain – or better: the central nervous system – is the seat of the mind and thereby of consciousness. Conducive to the success of this approach was the comparative study of pathologies and neurological anomalies, which demonstrates that certain mental functions are only possible if certain parts of the brain and their links are intact. While formerly neurologists and psychologists had to rely on posthumous autopsies to establish links between behaviour and brain structure, technological developments in recent decades have enabled increasingly accurate correlations. Today, devices such as PET scans or magnetic resonance scans provide the chance to follow changes in brain activity in patients as well as in healthy test persons, while molecular neurobiology experiments on animals test the function of concrete genes (Damasio 2000, p. 14).

To study pathologies in order to understand how the brain produces a sense of self might seem paradoxical but there are at least two good reasons for it. Firstly, in as complex organisms as vertebrates, correlations between lesions in the central nervous system and anomalies in perception or behaviour are often the best clue to understanding the relationship between structure and function. Secondly, anomalies in the sense of self can provide clues to the question if the self is indeed nothing more than a construct, an epiphenomenon of various physical functions that doesn't intrude in the causal nexus but is valued nevertheless.

Although it seems to be relatively easy for many people to imagine themselves as disembodied or as living in other bodies and other periods like Virginia Woolf's *Orlando*, this ease seems to be misleading. In fact, pathologies, which hinder a sense of being identical with or owning part of one's body are accompanied by a strong sense of discomfort and alienation. Patients who suffer from pathologies like asomatognosia fail to recognise parts of their bodies (Feinberg 2001, p. 8). They try to alleviate the unpleasant feeling which ensues by fantastic confabulations, arguing that the limb in question belongs to someone else or has a mind of its own.

One of the most famous and sensational examples of a pathology of the self is the Capgras syndrome (Feinberg 2001, p. 33). Patients suffering under it misidentify close friends or members of their family as imposters or aliens. While such patients are not impaired in recognising the person in question, they feel a lack of emotional familiarity with them. The ensuing dissonance leads them to fabricate stories to explain why this person they are faced with, although seemingly identical in every respect to their good friend or spouse, is in fact someone else. Once again, the lengths that patients go to in order to explain away the dissociation between their emotional response to a person and their cognitive recognition, indicates a preference for a unified self.

Further examples of pathologies of the self are autoscopia, in which the patient projects a copy of herself into the outside world (Feinberg 2001, p. 80), cases of depersonalisation, in which the patient feels a strong sense of alienation towards herself and her surroundings and which can take the form of a Cotard delusion in especially strong cases. Patients suffering from this pathology become so detached from their phenomenological experiences, that they can doubt being alive (Fine 2005, p. 49).

Possibly the most spectacular pathology of the self concerned with singularity and unity is the so-called “Multiple Personality Disorder” (MPD) or “Dissociative Identity Disorder” (DID) as it has become in recent years. A fierce debate rages over the phenomenon of dissociative identities or multiple personalities and there are some psychiatrists who reject the idea that it is a genuine affliction and argue that it is produced by therapy (Hacking 1995). DID is typically associated with early childhood traumas connected with abuse. It involves at least two but generally more alter egos, who alternate in taking control, amnesia and a lack of substance abuse or of a general medical condition (Hacking 1995, p. 19). DID patients often have a history of other mental ailments, most typically severe depression (Hacking 1995, p. 25). Such other ailments seem to be the initial reason of beginning therapy, with many patients being unaware of alter egos and denying their existence (ibid.). Some therapists therefore attempt to create co-consciousness of the alter egos to induce a method of peaceful co-existence in as integrated a manner as possible. Once again, the phenomenon of DID could be argued to indicate that disintegration or dissociation of the self does not only practically undermine effective agency but is highly unpleasant for the patient.

The Shifty Self

Confabulations are very common in such pathologies of the self but they are only one among the many tricks the self plays on itself during its self-creation. Other functions of the mind also serve to build and maintain a picture of the self – for itself. Among these is memory, which has long been counted among one of the necessary conditions for self. This does not simply involve the self-serving taints that the

memory of former selves acquire in time (and which must be one of the reasons, beside an urge to code one's life as having been a success, of the tendency to distort the memories of one's past, especially one's childhood and youth, the so-called fading affect bias (Fine 2005, p. 43). The memory of former selves as a narrative can be fairly rigid and has only an indirect effect on the present. There is another function of memory however, which pertains more to the self-representation in the present and this, too, is prone to self-serving distortions. Whatever are the memories that fit best into the present need of the self to appear in the best light are preferred by the present self and used to create it (Fine 2005, p. 12). This manner of speaking is misleading, however, since it implies that the same self is schizophrenically divided into subject and object and self-consciously creates itself out of the most flattering memory scraps.

The notoriously fickle memory is not the only tool the self uses to enhance itself in its own view. Rational judgement is another versatile helper in making self-serving judgements and choosing actions, while it makes use of memory to distort the facts and justify the choice already made (Fine 2005, p. 26). Also, depending on whether the own self is in question or someone else's, cognitive qualities, which are thought to be the tools for unbiased judgement and rational choice of action prove to be tainted by self-serving contortions of post-hoc rationalisations (Haidt 2001). It seems that among the information available, what is regarded as more relevant and more conclusive is significantly more often that which is more flattering or conversely, less threatening both to self and to life-choices. The figure of the rational, self-knowing agent is, it seems, less common and more pathological in "the real world":

There is in fact a category of people who get unusually close to the truth about themselves and the world. Their self-perceptions are more balanced, they

assign responsibility for success and failure more evenhandedly, and their predictions for the future are more realistic. These people are living testimony to the dangers of self-knowledge. They are the clinically depressed (Fine 2005, p.23).

These findings might imply that being too clear-minded might actually be detrimental for effective (moral) agency.

Empathy

Pathologies of the self, if the term is appropriate, help us to understand how a sense of self, both minimal and autobiographical, is created in an individual. They indicate that the seemingly given sense of self is nourished by a complex interaction of sources; neurological, psychological, social and – it might be argued – self-constitutional. Empathy is a further promising area of study because it is directly concerned with elements of relational ethics: Firstly it involves the idea that the mark of consciousness and thereby of the minimal, core self is intentionality. Secondly, in the case of empathy, the direction that intentionality takes is the consciousness of others. And thirdly, along with these two characteristics, empathy appears to be at the centre of relational ethics since it combines its defining characteristics: It is directed outward, i.e., it is other-related and it takes into account the morally relevant states of consciousness of these others, those that involve interests and/or rights.

A neurological link to empathy was claimed with the discovery of mirror neurons. The existence of mirror neurons was postulated when it was observed that the ventral premotor area of the frontal lobes of macaque monkey contains some cells, which fire when a certain actions involving the hand and the mouth were made. The interesting thing about these neurons was that they also fire when the individual

is simply observing instead of initiating such actions. This observation does not have to be of a visual kind; mirror neurons also fire when the subject hears the sound of an action (Iacoboni et al. 2005, p. 0529) Mirror neurons were hailed as the most important discovery of the decade, similar in its importance to the discovery of DNA and promising insight into empathy, learning by imitation or language. The link between these cells and empathy seems so strong that Vilayanur Ramachandran calls them “empathy neurons” or even “Dalai Lama neurons” that “dissolve the barrier between self and others” (Ramachandran 2006).

Being able to infer the beliefs and desires of others has obvious evolutionary advantages, not only and not even predominantly with the effect of enabling or strengthening moral agency but rather by endowing the owner of such capabilities with the power to develop a matching theory of mind and thereby with the possibility of predicting behaviour and selecting methods of influencing and manipulating it. Due to this skill human beings are called the “Machiavellian Primate” (Ramachandran 2000). How it is that we are able to know that others have minds is a well-known problem. How it is that we systematically attribute certain beliefs and desires to these other minds goes a step further. The question is much more than a fanciful past-time of philosophers; rather, it has vital implications for the status of moral patiency to be attributed to beings. The notorious insistence of Descartes, for example, that non-human beings were simple automatons that had no mental life and thus no sensations put them on the same level in the moral community as machines such as clocks.

Mirror neurons can be interpreted as providing an answer to the question of how we infer that other beings have minds in general and certain beliefs and desires in particular. An alternative interpretation is that their function is action recognition

(Iacoboni et al. 2005, p. 0529). Action recognition differs from the recognition of motor acts in that it involves the attribution of a goal and an agent. Thereby action recognition goes beyond the recognition of the movements involved in the action but includes the beliefs and desires of the observed individual. Reaching out and grasping an object for example is interpreted as “wanting an apple” (ibid.). The question regarding motor-neurons is then, how far they inform the individual about the intentions of the observed individual. Since it is their being embedded in contexts that informs about the intention of actions, the question is whether the same grasping action with or without context elicit the same activity in mirror neurons (ibid.). The researchers around Iacoboni monitored the brain activity of individuals watching videos of a tea set before usage, after usage, a cup in the tea set being grasped by a hand, a cup in the tea set being held by the fingers only and two pictures of a cup being grasped or being held by the fingers only without any context. Observing grasping actions in contexts were accompanied by greater activity in mirror neurons than those without context or those of context without action, which suggests that the mirror system in humans serves not only for the formal recognition of actions but “also constitutes a neural system for coding the intentions of others” (Iacoboni et al 2005, p. 0530). The increased activity of mirror neurons in the right inferior frontal cortex during the conditions set for “intention” had not better be interpreted as caused by the increase in the complexity of the condition, i.e., by an increase of objects because there were differences between the cleaning intention vs. drinking intention (Iacoboni et al 2005, p. 0533). In summary the authors argue that the “present findings strongly suggest that coding the intention associated with the actions of others is based on the activation of a neuronal chain formed by mirror neurons coding the observed motor act and by ‘logically related’ mirror neurons

coding the motor acts that are most likely to follow the observed one in a given context. To ascribe an intention is to infer a forthcoming new goal and this is an operation that the motor system does automatically” (ibid.).

Localising the Self

One challenge is to reconcile the idea and the feeling of a unified self with the neurological fact that there is no place in the brain where the innumerable goings-on converge to create such a self. A popular view is that the brain is hierarchically structured and that certain areas correspond to certain conscious or unconscious mental events, all with a specific role in the emergence of consciousness at the highest point of convergence. A modification of this architectural picture is a hierarchy that is not pyramidal but nested; here the self is understood as the product of the co-operation of parts of the brain (Feinberg 2001, p. 7).

A phenomenological approach to the question of the location of self leads to a similar identification of the brain with the mind. It is an open question whether it is the amount of sense organs collected in the head, the convergence of their afferent paths in the brain or the concentration of the parasympathetic nervous system in the head and the breast that makes the head and possibly the upper part of the torso feel like the seat of the self. There are moments when other parts of the body seem to be more prominent than usual; the sexual organs during sex or a limb in pain. But even in such cases, the self seems to be 'up here' with at most an extension of it radiating to the part in question, not a displacement. The sense of the location of self is further dynamic in that the boundaries of the body; the skull or the skin doesn't seem to hold it in. Self feels more like a field, an aura than an organ and its physical boundaries

feel beyond those of the head. Nevertheless, the amount of correspondence between the phenomenological and neurological attempts at locating the self is striking.

While the sense of self as located in space may correspond to the location of the brain and certain parts of the central nervous system, a thicker sense of identity involves the objects around us that are not part of the self in the former, locative sense; “The self is a continuum of relationships” (Feinberg 2001, p. 30). This continuum can reach from a feeling of a detached, pure ego to the other, who is considered as part of the self. Self is not a constant point in this continuum but dynamic, it can be nearer to the sense of pure, perceiving ego or further and its position in a given moment of time does not depend on itself alone but on the nexus of relationships and attitudes it finds itself in. “The self does not exist as a rigid structure, in the way our outer skin separates us from the world. Rather, like the amoeba, the self displays an uncanny ability to change its shape, alter its margins, reform and regenerate new parts as needed” (Feinberg 2001, p. 50). Near-pathological manifestations of dysfunctions in the balance between the emptiness of the idea of the pure ego and the diffusion of the interrelation with (biographically) distant objects include feelings of *jamais vu* and *déjà vu* (Feinberg 2001, p. 31). One significant point about these phenomena is that they feel uncomfortable; a further indication that a balanced, unified sense of self both in relation and in contrast to the world is preferred.

An important disclaimer at this point is linked to many traditions, which consider the sense of a singular or unified self as a hindrance to the achievement of some defined good. They advocate a purposeful striving to overcome the feeling of a unified self. This advocacy may have two reasons: The sense of self as singular or unified can be thought of as an illusion, which blocks the access of truth. Such

conceptions of self are familiar from Eastern philosophical traditions as well as Schopenhauer's veil of Maya and his related advocacy of quietism as an aspect of *ars vivendi*. The other reason is also based on the idea that the good life requires a destruction of the sense of a singular or unified self could lead to a life lived more fully.

Returning to the material localisation of self: If we take the urge to confabulate as a sign of an effort to overcome cognitive dissonance or a need to (re-) establish a „narrative centre of gravity“ (Dennett), then the fact that fantastic and spontaneous confabulations can be observed most often in patients with damages to the frontal lobes (Feinberg 2001, p. 68) might serve as an indicator that the frontal lobes contribute to the coherence and narrative integrity the self in the healthy individual. However, there are cases of patients whose frontal lobes have been removed and who nevertheless maintain a sense of a unified self (Feinberg 2001, p. 106). Pathologies, it seems, can do no more than indicate which parts of the brain are involved in creating the self. The whole process seems to be too complex to be fully understood yet.

A further candidate area for the seat of the self is the language centres in the left brain, especially for adherents of the view that beside memory, language plays a central role in the production of self. More generally the left brain seems to have a major role in creating the feeling of a unified perceptive self by unifying incoming perceptions (Taylor 2008). The unity of vision is a good example of the unity of consciousness in general: The mind succeeds in making one unified sense of vision out of two, slightly different perspectives on the same view. The „Mind's eye“, the product of the two is created as if it were a single eye between the two eyes (Feinberg 2001, p. 111). Sight is produced in a classically hierarchical, bottom-down

build-up: Each cell in the retina has its own receptive field and these converge in the area V1 (Feinberg 2001, p. 112). The higher in the hierarchy, the more specialised the cells become; so much so that some higher order cells respond best not only to faces but some to faces from the frontal view while others prefer the profile (Feinberg 2001, p. 113). This phenomenon has led to the joke that there might be a cell so specific that it responds only to the face of one's grandmother. How such numerous and diverse specialised bits of information converge to create a unified sensory view of the world is the so-called “binding problem”: How does the brain create a unified perceiving subject, which binds the sensory inputs from the sensual organs and does recognise the grandmother as sitting in a chair in a red dress with the green background of a garden and smelling of roses? This question is related to – but not identical with – the question of how consciousness is created from the simultaneous firing of all the cells participating in the creation of that singular situation. The same is the case for the sense of a unified self as the initiator of action. The question of how various sensory inputs are unified to create the consciousness of a given moment is as yet as unanswerable as the question of how the “will” to initiate an action seems to cause the relevant changes in the body and in the world as part of the Mind-Body Problem.²²

What should be noted is the affinity of the nested hierarchy view of numerous areas in the brain, converging on one singular sensation resembles the hierarchy of monads as Leibniz envisioned them. At a given moment there are infinite numbers of monads at work with some having very elementary apperceptive functions while others, like the cells recognising faces in the profile also having apperceptive

²² It is remarkable how the terminology used in this problem follows the familiar pattern of apartheid between humans and non-human animals; the challenge of explaining of how the „top“ of the hierarchy succeeds in initiating 'behaviour' – in non-human animals this is called 'instinct', in humans its 'will'. (Feinberg 2001, p. 118).

functions up to perceptive monads that recognise the face consciously. Their totality perceives the lines, the curves, the smell, the colours that make up the grandmother and finally the emerging consciousness of the constellation of monads: “The brain creates the unity of the self by producing a nested hierarchy of meaning and purpose, where the levels of the self, and the many parts of the brain that contribute to the self, are nested within all other levels of the hierarchy” (Feinberg 2001, 149).

Nevertheless, the search for the self in the brain by trying to locate one specific centre where 'all comes together' is bound to fail. There is no Cartesian pineal gland to solve the problem of the unity of consciousness, of how the body and the mind interact or of how consciousness emerges from the structure and activities of a vast number of connected nerve cells. The brain can be thought of as analogous to the body, with various areas carrying out various functions, which are more or less necessary for the functioning of the whole. In a similar way, there are numerous centres in the brain, which need to co-operate for sensual perception while other functions seem to be localised in a few very specific areas (e.g., Wernicke and Broca areals in the left brain for language). The function of most of these areas have been extensively mapped out, the most famous among them being the homunculus, which is drawn from the body's sensory map and the analogous motor control systems in the cortex (O'Shea 2005, p. 59). Numerous parts of the brain specialise on certain functions that normally create a singular sensory experience: “Where it is a question of 'mind' the nervous system does not integrate itself by centralisation upon one pontifical cell. Rather it elaborates a million fold democracy whose each unit is a cell” (Sherrington, 1947 quoted in Feinberg 2001, p. 111). The co-operation between the units and the whole should not be understood as an arboreal hierarchy but rather as a 'nested' one in which the units are dependent on and constrained by each other,

with no clear centre of control: “In the same way that mitochondria and the lung contribute to the life of the person, in the nested hierarchy of the mind, all the lower order elements - every line, shape, and patch of color that make up our awareness of the face continue to make a contribution to consciousness” (Feinberg 2001, p. 130).

Such pictures of the fairly detailed and accurate map of the functional centres in the brain and their interaction, the genesis of sensations and the production of motor action is in a stark contrast to the “hard” questions concerning the brain, the mind, consciousness and the self, which are as unanswerable as they have ever been. Two questions are concerned with the production of self: Firstly, how is it that the afferent and efferent nerve cells interact to produce a feeling of a unitary self? This is mainly a neurological question and closely related to the not-so-hard question of consciousness. There is reason to hope that there might be an answer to this question in the future (Searle 1984). The second question is related to the hard questions concerning consciousness, free will and agency. These are not so much physical but philosophical question and there is so much confusion about the right way to pose the question – let alone about answering it – that it is not as easy to be hopeful about a solution any time soon.

The anatomy of the brain further leads to the phenomenon of split-brain patients, popular in the discussions of multiple selves and consciousness. The neocortex is divided into two hemispheres, which are connected by a pathway of neurons called the corpus callosum, consisting of about a million axons with an equal amount of neurons from both sides (O'Shea 2005, p. 58). If the corpus callosum is intact, the two hemispheres together produce “a single seamless perception of the world” (Taylor 2008, p. 16). If the corpus callosum is severed – due to injury or as it was customary to do in cases of severe epilepsy – then the perception of the world

and the orchestration of motion continues to be unified in most cases and at most times because the corpus callosum is not the only way by which the two hemispheres communicate and co-ordinate. Such communication can be obstructed in experimental settings and these, together with studies of patients with pathologies targeting large parts of one hemisphere provide clues as to functional centres in the brain on the one hand and the method of co-operation between these centres on the other.

In the division of labour in the brain, motor control and sensory representation of one half of the body is most often undertaken contralaterally, i.e., by the opposite half of the brain. Sensory data of the right side of the body ends up in the left hemisphere and vice versa while the efferent nerve cells controlling motor action of the right side stem from the left hemisphere and vice versa. In sight, the left halves of the retinas send their impulses to the left hemisphere and the right halves to the right hemisphere, while smell is transmitted ipsilaterally from each nostril (Nagel 1971, p. 149). There is a limited amount of ipsilateral sensory and motor control as well, especially of such parts of the body that are in the middle (Feinberg 2001, p. 90). The brain is also “split” when it comes to specialised functions, most prominent among these being the language centres that are located in the left hemisphere. Attentional control or certain aspects about emotional behaviour are among the functions of the right hemisphere (Feinberg 2001, p. 91).

Studies on split brain patients have shown that in certain experimental settings, which provide one hemisphere with information that the other lacks or which provide conflicting information, the result can be that the hemispheres present incompatible or conflicting behaviour. This indicates that the sense of unity in perception and in co-ordinated behaviour is something that the two hemispheres must

have learned to do (Feinberg 2001, p. 94), not a given. Such studies are not absolutely conclusive. It seems that depending on whether the corpus callosum is intact or not, the two hemispheres act differently; i.e., when connected they are more co-operative, enhancing and complementing each other while the dysfunctions that appear in split-brain patients are not necessarily what is “uncovered” by the surgery but what is produced by it (Taylor 2008, p. 28)

Nevertheless it seems indubitable that there are certain functional differences between the two hemispheres: The right hemisphere, for instance, is responsible for creating a “master collage” (Taylor 2008, p. 29) of the incoming sensory information and is concentrated on the present. Some of such functional differences between the hemispheres have direct implications for morality: “Our ability to be emphatic, to walk in the shoes of another and feel their feelings, is a product of our right frontal cortex” (Taylor 2008, p. 31). The left hemisphere on the other hand organises experience in a linear succession, it constructs a linear narrative out of the uncoordinated details of the right hemisphere. Since the language centres are in the left hemisphere, the left hemisphere is thereby the one which creates thought coded in language as opposed to the more direct perception of the right brain. Both functions – bringing moments into temporal order and language – make “cogito” a product of the left brain. Additionally, it is the left hemisphere that creates habitualised patterns of response to certain clusters of stimuli, which are used in automatised action and heuristics: “From a neurological standpoint, every time a circuit of neurons is stimulated, it takes less external stimulation for that particular circuit to run” (Taylor 2008, p. 32).

The phenomenon of split brain patients is often used to argue that there is not one unified consciousness per brain but (at least) two. As early as 1780 it was

claimed that humans were *Homo Duplex* (Taylor 2008, p. 27). Other syndromes indicate that there might actually be as many “selves” or at least centres of perception and decision-making as there are functionally organised centres in the brain. One example that indicates that this might be the case is the syndrome of the alien hand. An alien hand does not only feel alien, as in the case of patients who suffered from asomatognosia, it also seems to have a “mind of its own” and may act contrary to the patients will or even hostile to the patient, trying to hinder or even harm them (Feinberg 2001, p. 94).

Nevertheless, what is most interesting about patients with split brains, with asomatognosia or with an alien hand is the fact that in most cases, their experienced sense of self remains unified despite everything. One way to explain this is the fact that even when the corpus callosum is severed, there is enough ipsilateral projection of information through the lower levels – the spinal cord and the brainstem – of the brain (Feinberg 2001, p. 99). Beside these lower levels there is another pathway, the anterior commissure which ensures some transfer of information between the two hemispheres (Feinberg 2001, p. 100). Such pathways explain how it is possible that both hemispheres receive information from both ipsilateral and contralateral sensual sources or are able to organise co-ordinated motor action. This would indicate a system of an imperfect but functioning bottom-up information flow. More interesting are cases where such information is not available through any pathway but where the patients experience is unified and coherent nevertheless. In these cases, patients make use of “confabulatory completion” (Feinberg 2001, p. 101), where they depend on their former experience of the world to fill in the blanks and invent a coherent subjective sense of themselves and the situation they find themselves in. Being experienced with how the world works, this type of confabulation is an important

heuristic that helps patients deal with the world without it being obvious to them or to outsiders that they are actually confabulating. These cases seem to indicate that while there are indeed at least two centres of consciousness and volition in the brain, which can have independent or even contrary experiences and patterns of behaviour, it seems that experiencing unison is a higher order interest that both – or all – parties are willing to co-operate in to create.

CHAPTER 4

MARTIN BUBER'S MORAL ONTOLOGY

General Introduction

It is not easy to situate Martin Buber in the philosophical landscape, partly because he avowedly defies categorisations and rigid objectifications. He defines himself as an “atypical person” and assumes that his aversion towards “the common excessive typology ultimately originates in this fact” (Buber 1961, 1111, translation B.S.). He rejects categories that others have believed him to espouse and he therefore cannot be introduced easily with such terms as life philosophy or existentialism. The nearest one can get to a denomination, which he might accept is probably his brand of a philosophical anthropology; “For his concern is with man, with man not only in his disparate interests, but also in his wholeness – in the possibility open to him of wholeness of being” (Smith 1967, p. 10).

Buber's ontology of the self and his moral philosophy are deeply intertwined. So much so that he himself interprets his discussion of good and evil as “a contribution of an ontological ethics” (Buber 1953, Introduction). Comparably, “[t]he dialogical principle is an ontological one because it is concerned with a basic relationship between man and being; hence with the being of man, since this is grounded in his relationship to being” (Buber 1964, p. 22).²³ He acknowledges the great variety in moral judgement and the influence of social convention on human

²³ The translators of Buber into English tend to use the word “man” where Buber writes “Mensch”, i.e., human being.

agency. His conviction that nevertheless there is a realm of the moral is an ontologically realist position:

In spite of all the problematic of moral judgement, in spite of the constitutive impermanence of moral valuation, we must recognise and accept that in human reality there does indeed exist a specific of this kind, a specific, not according to valuation and judgement, but in being itself, and that this specificity is evidenced precisely in the fact that there things happen differently than otherwise in the life and the soul of man. It would, therefore, be totally insufficient to refer the matter to the existence of states whose nature and course are influenced by the 'moral censorship' of society, whether this censorship is the cause of submission or of rebellion; there can be no question at all here of the psychology of 'inhibitions and 'repressions', which operate no less against some social convention or other [...] (Buber 1953, p. 116).

As we shall see in more detail in the discussion of the two possible modes of being of humans, it is the "double relation" that humans have to being, rather than being itself, that is the ontological interest of Buber (Buber 1963, p. 592)

Another term with which Buber has come to be associated and which he himself has emphasised a number of times as central in his writings is that of "dialogue". Hilary Putnam²⁴ discusses three philosophers of dialogue, Rosenzweig, Buber and Levinas (Putnam 2008) and draws parallels between the pragmatist position and his own, when he rejects both moral realism and moral subjectivism "[...] like the classical pragmatists, I do not see reality as morally indifferent: reality, as Dewey saw, makes demands on us. Values may be created by human beings and human cultures, but I see them as made in response to demands that we do not create. It is reality that determines whether our responses are adequate or inadequate" (Putnam 2008, p. 6). That reality is the reality of relationships; it's humans involved

²⁴ Putnam himself is an example for the parallel existence of two incompatible approaches to the world. He is aware of the gulf between his faith on the one side and his physicalistic-materialistic outlook on the other: "Those who know my writings from that period [philosophy of science in the earlier part of his career, B.S.] may wonder how I reconciled my religious streak, which existed to some extent even back then and my general scientific materialistic worldview at that time. The answer is that I didn't reconcile them. I was a thoroughgoing atheist and I was a believer. I simply kept these parts of myself separate" (Putnam 2008, p. 4)

with others – other things, other beings, other humans – in relationships (Putnam 2008, p. 58). Putnam's position incorporates three main elements of dialogical philosophy in general and Buber's version in particular: the demand that the situation makes on the moral agent, dialogue with the other in the situation and relationship as the modus of being.

Relationships are central for Buber's thought but not in the sense that a person who, like a monk, chooses to isolate herself and keep her own company would lose out on some essential aspects of the good life. His position is more radical in that he believes that human beings are essentially relational beings in an ontological sense. The other and the relationship one has with the other do not only promote a good, they are part of a couple, who are ontologically dependent on each other. For Buber, the subject in itself does not exist, it is ontologically dependent on the object:

The doctrine of immersion demands and promises penetration into the thinking One, 'that by which this world is thought', the pure subject. But in lived actuality no one thinks without something being thought; rather is that which thinks as dependent on that which is thought as vice versa. A subject that annuls the object to rise above it annuls its own actuality (Buber 1970, p. 137).

These elements are demonstrated in a biographical anecdote of Buber as a child, when he lived with his grandparents on their farm after his mother left him and his father:

This precocious and undoubtedly very isolated and lonely little boy devised for himself dual-language conversations between a German and a Frenchman, later between a Hebrew and an ancient German. Through these conversations he came 'half in play and yet at times with beating heart', to feel the tension between what was heard by the one person thinking in one language and what was heard by the other person thinking in another. [...] Indeed, it is here that one can find the essence of dialogical understanding not as precise definition, technical communication or subjective empathy, but as 'inclusion' (*Umfassung*) – experiencing the other side of the relationship while not losing the awareness of one's own and of the polar tension between one's own and the other (Friedman 1981, 7).

The inclusion in question is fundamental for Buber's brand of ethics. Although he sometimes creates the impression that his writings deal with metaphysics, mysticism or religion rather than ethics, his concern is always ethical or always also ethical. As Buber understands it, living authentically and rightly involves the capacity to be aware to the demands of the situation and to be able to answer to it. The answer is "untranslatable", as is the demand and should be understood as the language of action and omission (Buber 1930, p. 163). As such, Buber's outlook on life is always also ethical in the narrowest sense of the term that seeks answers to questions of what to do. Dialogue and inclusion are not simply means to that end but essential elements of the morally good way to live.

Considering his methodological approach Buber is again something of a rarity as he explicitly opposes the idea of a systematic account of that what he is trying to convey. He does not attempt to develop a doctrine (*Lehre*) as he writes in a much quoted passage.²⁵ Rather, he wants to be understood as being in a dialogue with the reader: "When it comes down to it I appeal to the real and possible life of my reader. The intention of my writings is really a quite intimate dialogical one" (Buber quoted in Reichert 1996, p. 287, translation B.S.). His starting point is experience and he appeals to the experience of his readers, without reducing such experience to subjectivity (Buber 1963, p. 592). It is because of his conviction that the sort of experience, which he regards as foundational for his philosophical anthropology is "accessible to all in some measure" (Buber 1964, p. 18) that he appeals to it.

²⁵ „I may not go beyond my experience and never wanted to be able to. I stand witness for experience and I appeal to experience. [...] I have to repeat it again and again: I don't have a doctrine (*Lehre*). I only show something. I show reality, I show something about reality, what hasn't been seen or has been seen too little. I take the one, who listens to me by the hand and lead him to the window. I open the window and point outside. I do not have a doctrine but I hold a conversation" (Buber 1963, 593).

Buber repeatedly remarks on the fact that his contributions to philosophical anthropology or anthropological philosophy do not follow a method. Nevertheless, there are a few remarks that give a more positive idea of his methodology.

Comparing psychological introspection and retrospection for example, Buber shows himself to be sceptical about the former, following the more and more prevailing post-Freudian criticism of his day:

A few modern psychologists have indeed recognized that so-called self-observation exercises transforming influence upon the psychic process comparable to that which the physicist has posited for the observations of electrons. Therefore the retrospective method that works with the more or less reliable events of an unarbitrary memory is the more useful of the two (Buber 1964, p.59).

It is not clear whether Buber means the individual memory or objective records of the individual's biography, since it is doubtful whether the notoriously fickle memory can be regarded as more reliable. More clear is his distinction between psychological and anthropological methods, for while psychology reduces events to their "psychic side", the anthropological method, as Buber sees it, is "integrative". Buber argues for the superiority of the anthropological method over the psychological one since the former approaches the human being as a more complicated being and does not ignore embodiment (Buber 1964, p. 59). Some further remarks propose an anthropological method that is reflexive in that it does not ignore the fact that the object of the investigation is at the same time its subject:²⁶

Here, where the subject is man in his wholeness, the investigator cannot content himself, as in anthropology as an individual science, with considering man as another part of nature and with ignoring the fact that he, the investigator, is himself a man and experiences his humanity in his inner experience in a way that he simply cannot experience any part of nature – not only in a quite different perspective but also in a quite different dimension of being, in a dimension in which he experiences only this part of all the parts of nature (Buber 1938, 147).

²⁶ A generation after Buber, Pierre Bourdieu similarly demanded a reflexive approach in anthropology and sociology (Bourdieu, Waquant 1996).

This double access to the human being, i.e., the first personal and the third personal makes a reflective anthropology necessary but it additionally grounds the necessity for a phenomenological account of the human life. Approaches that emulate the objectivity of hard sciences are bound to fail, not only because they try to avoid the problems that first-personal accounts make, by ignoring it. Such a dismissal of the first-person account simply causes the account to lack half. The philosophical anthropologist must incorporate his subjectivity into his account of the human (Buber 1938, p. 148).

The second methodological element concerns Buber's relational ontology. In his exploration of Kant's formulation of the anthropological question concerning what man is, Buber comes to the conclusion that "an individualistic anthropology, an anthropology which is substantially concerned only with the relation of the human person to himself, with the relation within this person between the spirit and its instincts, and so on, cannot lead to a knowledge of man's being [*Wesen*]" (Buber 1948, p. 236). Human beings' *Wesen* is relational and it can take two different forms according to the type of relation they find themselves in. Philosophical anthropology must therefore not only be reflective, i.e., not only must it account for the fact that the study of the human being must take the first person view into account. It must also consider the human being as always in a relation and must therefore consider the nature of the relations and the partners.

The place that Buber accords to reason is a limited one. The reason for this limitation is not an underestimation of the worth of reason in (moral) agency but rather the postulation that rationality has been conceded a greater amount of authority than is due to it. As such, Buber's critique of modernity and of reason is

typical of the critique of enlightenment, characteristic of the first half of the twentieth century. As one of the elements of experience rather than in its “detached, egomaniacal form”, reason contributes to the evaluation and handling of a situation, which is “with necessity a philosophical one, and this means: a logicalising one; everything depends on the indispensable ability to reason not to misjudging its post and not behaving like the proper addressee.” (Buber 1964, p. 1112). Reason ought to be responsible for enabling consistency but it ought not to “sacrifice anything of that reality, which the experience demands” (Buber 1964, p. 1112). This will not lead to a system as is common in philosophy but to what Buber calls a “consistent, conveyable structure of thought” (Buber 1964, p. 1112). All in all and despite his frequent disclaimers, Buber’s approach seems to contain a specific methodology; one that makes use of reason, consistency and conveyability, but not verifiability or falsifiability.

Buber’s insistent rejection of building a philosophical system or a hierarchy of principles is partly based in the importance that he accords the situation. Every situation, every agent is unique as is the constellation of the two and barely anything can be postulated beforehand about either of them. One further problem is related to the terminology that is specific to Buber and his seeming tendency to use them in a variety of connotations. The danger of misunderstanding a certain passage is high if the reader is not prepared to demonstrate the same alertness, attentiveness and open-mindedness that Buber expects from the moral agent. These aspects are among the difficulties in reading and writing about Buber. Wolfgang Krone questions if, since Buber is explicit about his rejection of a systematical methodology, the attempt to develop a systematic account of his work is not predestined to fail (Krone 1993, p. 37). He argues that Buber cannot be understood by a traditional methodology since

his intention is not to persuade the reader but to invite her to a dialogue, in which she is expected to turn to her own experiences.

In an interview with Walter Kaufman, Buber concedes that the traditional (pre-) occupation with concepts is “a central task of thought because it is the presupposition for an ever-renewed confrontation with reality” (Buber 1964, p. 17). Nevertheless, he regards this occupation as merely a means to an end, which in itself is not philosophical or rather cannot be “grasped philosophically” (ibid.). The clarification of concepts is an important tool but it should be approached critically and not taken to be the only or the main aim of philosophy. In Buber’s preferred articulation using metaphors: “Socrates overvalued the significance of abstract concepts in comparison with concrete individual experiences. General concepts are the most important stays and supports, but Socrates treated them as if they were more important than bones – that they are not” (Buber 1964, p. 67).

The Moral Situation

Although I use the term “moral situation” to describe Buber’s situationism, I only do so to emphasise that for Buber the situation is the proper domain of the moral and the everyday life is the proper domain of the situation. He himself does not specify or emphasise the situation as a moral one. His brand of moral philosophy is resolutely against what he terms *Sonderethik*: “The idea of responsibility is to be brought back from the province of specialized ethics, of an ‘ought’ that swings free in the air, into that of lived life” (Buber 2002, p. 18). Ronald Gregor Smith has translated the term as “specialized ethics” (Buber 2002, p. 18) but this translation might miss the point a little. What Buber criticises here as *Sonderethik* is not the idea

of specialisation as expertise but the idea that there is a certain, separate area that is concerned with ethics and in which it is appropriate and necessary to be concerned with ethics. This area would remain singular and separate from daily life. Daily life for Buber, however, is nothing but the proper area for ethics and – in his religious streaks – of worship. One of his main dictums is that both should be brought back to the earth and to singular, particular situations in daily life:

But the sounds of which the speech consists – I repeat it in order to remove the misunderstanding, which is perhaps still possible, that I referred to something extraordinary and larger than life – are the events of the personal everyday life. In them, as they now are, ‘great’ or ‘small’, we are addressed, and those which count as great, yield no greater signs than the others. (Buber 1930, 162 / Buber 2002, 19).

The situation is central for Buber. It is there that the agent has to prove herself by remaining alert and aware of the demands of the situation and able to accept the responsibility that such responsiveness brings with it. Buber tends to use the metaphors of speech, listening and answering when describing the morality of the situation. The German word for responsibility, as one of the central demands that Buber puts on the moral agent is *Verantwortung* and is a derivate of the word *Antwort*; i.e., answer. The answerer is alone with her responsibility; there is neither system nor a set of rules of the community she belongs to, to which she can turn for ready-made answers; “Maxims command only the third person, the each and the none” (Buber 1939a, p. 136). The state of being without a wholly reliable hold and the loneliness of the agent in the moral situation is something of an existential trait in Buber’s moral theory.²⁷ Nevertheless, although she will never be able to respond in

²⁷ Buber, whose aim is to resist the temptation to categorise phenomena and build structures of thought and explanation, naturally protests against being labelled with any –ism; “But if those be called existentialists who transpose human existence itself into the center of rational contemplation, then one could call me that” (Buber 1964, p. 18).

the ideal way, she will respond in a way that is in the “right way” or appropriate (*rechtmäßig*):

I give the word of my answer by accomplishing among the actions possible that which seems to my devoted insight the right one. With my choice and decision and action – committing or omitting, acting or persevering – I answer the word, however inadequately, yet properly; I answer for my hour. My group cannot relieve me of my responsibility, I must not let it relieve me of it [...] (Buber 2002, 80).

Since attentiveness²⁸ and the capability of perceiving the demand of the situation are the main qualities needed to be a moral agent, who or what hinders someone from this perception, for example by giving an answer,²⁹ keeps the moral situation from evolving in the best possible way. Conversely a decline in the moral success of a person is directly related to the decrease in the capability of being open to a singular, particular and unprepossessed approach to a given situation. Buber therefore chooses open-mindedness as the most valuable thing that humans possess (Krone 1993, p. 116). These are to be employed in everyday situations, as many of the Chassidic tales³⁰ quoted by Buber aim to convey; the proper response of someone who has to deal with, for example dirty pots and pans is to immerse oneself, highly concentrated and with a unification of the self on that particular chore (quoted in Reichert 1996, p. 270), otherwise they will be accused of being “patchwork” (Reichert 1996, p. 173).

The realm of the ethical is therefore the everyday life and more particularly in particular, unique situations:

Life does not take place in that I play the mysterious board game with myself but in that I am faced with the presence of being, with which I have not agreed

²⁸ Sorabji also suggests that the „most useful of the ancient suggestions, I think, was that we need to postulate a faculty of attention” (Sorabji 2006, p. 11). A neurological link to attention exists in the ascending reticular activating system (ARAS), which controls arousal and “primes the cortex for stimulus reception, whereas the polymodal association cortex controls and focuses this arousal energy for attention” (Mendez & Gershfield 2004, p. 32)

²⁹ The word that Buber uses here is “*Einsagen*”; the act of someone other than respondent secretly giving her the answer, i.e., a common practice among schoolchildren.

³⁰ Buber does not aim to provide a “historically or hermeneutically comprehensive portrayal of Chassidism” (Buber 1963, 627, translation B.S.)

upon any rules of the game and with which it is not possible to agree on any. The presence of being, before which I am put, changes its shape, its appearance, its disclosure, it is other than me, often frighteningly other, and other than I have expected it to be, often frighteningly so. If I withstand, if I move towards it, if I encounter it genuinely, i.e., with the verity of my whole *Wesen* (being), then, and only then am I 'actually' there: I am there when I am there, and where this 'there' is to be, that is determined not so much by me but by the presence of being, which alters in shape and appearance. If I am not really there, then I am guilty. [...] The original state of being guilty is the Remaining-by-oneself (Buber 1948, 363, translation B.S.).

In other words *Situationslosigkeit*, the state of a lack of situation, which can be brought upon by fleeing the situation or by transforming it, be it with drugs or by self-deception (Friedman 1963, p. 164). It is necessary to be alert to the demands of the situations:

The situations have something to say! And the real, the biographical or historically real situations are not simple and flat like principles, they carry the contradiction in them, they lift it up to our face, and we may not ignore it, because reality lies in contradiction. 'All or nothing' does not hold, what holds is to realise as much as possible from our truth, as the impartial advancing insight in all the contradictions of the situation allows (Buber 1963, 618, translation B.S.)

The emphasis that Buber puts on the importance of open-mindedness and attentiveness is relevant to our discussion of situationism versus globalism. Buber believes that the judgements of agents tend to get encrusted (Krone 1993, p. 116), i.e., the older an agent gets, the higher the probability that she will judge and act in a more habitualised manner and in accordance to how she has acted in similar situations before. Critics of globalism like John Doris would object to this claim based on studies that seem to demonstrate that agents behave similarly in similar situations, indicating that the situation rather than the character of the agent has a high influence on behaviour. Supporters of the view that practical consistency is one of the main criteria of morality would object to Buber's claim that the progressive uniformity of moral responses is a thing to be deplored.

Buber is wary of absolutes, of all-or-nothing stances or of clear and unyielding positions. Therefore what I termed his existential trait only takes us so far. It is true that according to Buber we never know what to do before we are faced with the situation and we never know if we acted correctly afterwards: “Emotions about what the absolute force that one believes in – God, moral truths, conscience – may mislead the agent, there is no certainty that the choice made was the best possible one even if it is accompanied by a feeling of conviction that one has hit the mark (Buber 1936, p. 246). The agent can never know if she “is addressed by the absolute or by one of its monkeys” (Buber 1953a, p. 592). Such ready-found systems might not only mislead, they might also lull the agent into a false sense of security, which is “worse than real desperation” (quoted in Reichert 1996, p. 170). Thinking that one knows what method of moral judgement to use involves the danger of agents losing the sense of encounter and uniqueness that a situation brings, with no certainty of how to respond best. The picture Buber draws does not correspond to some lonely Sisyphus-like figure faced with a meaningless task however

I do not in the least mean that a man must fetch the answer alone and unadvised out of his breast. Nothing of the sort is meant; how should the direction of those at the head of my group not essentially enter into the substance out of which the decision is smelted? But the direction must not be substituted for the decision; no substitute is accepted (Buber 1936, p. 246 / Buber 2002, p. 81).

The moral situation is one about which little can be said beforehand. There is an exception however, when it comes to the agent who believes herself to have acted wrongly. Buber’s discussion of such cases of guilt and senses of guilt differs from what can be called the regular moral situation (Buber 1957). For one thing, the approach that Buber recommends for the failed moral agent involves reflection on the misdeed and as such the part that moral cognition is to play is much bigger. A second step is an awareness of diachronical identity; the failed agent is asked to bear

in mind that the repenting person is the same person who has failed in the past. This aspect is related to more traditional views regarding identity and responsibility over time. Lastly, the moral agent is expected to consider retribution through active commitment to the world (Buber 1957, p. 490).

One problem that Buber sees in the tendency of attempting grand theories and systems that explain the world is that they work via sterilisation (Buber 1930, p. 154). Sterilisation is the disavowal of the address; it is regarding the world in the way hard sciences do as the totality of entities and events that react according to discernible laws. In Buber's anthropological view the human being is not in an unrelated opposition to the world but is an addressee: "What occurs to me addresses me. In what occurs to me the world-happening addresses me. Only by sterilizing it, removing the seed of address from it, can I take what occurs to me as a part of the world-happening which does not refer to me" (Buber 2002, p. 13). This passage is open to misconstructions as Buber is fully aware. It evokes the idea of a superstitious belief in a universe of meaning beyond what physical, biological – or, as Buber adds, sociological – investigation allows. Buber argues that the key point of differentiation is not whether the world can be explained by scientific methods or should be understood as having a hidden meaning that can be enclosed by augurs. He considers these to be, on the contrary, as comparable in that they both regard the world as something that provides signs that can be read and interpreted. Both believe in the possibility of "looking up": "The common signature of all this business is that it is for all time: things remain the same, they are discovered once for all, rules, laws and analogical conclusions may be employed throughout" (Buber 2002, p. 14).³¹

³¹ This is the negative description of what Buber is trying to delineate. For a more positive one it is worth to quote a passage in full: "Real faith – if I may so term presenting ourselves and perceiving – begins when the dictionary is put down, when you are done with it. What occurs to me says something to me, but what it says to me cannot be revealed by any esoteric information; for it has never been said

In one sense it is a fairly straightforward demand that Buber makes when he insists on the importance of the situation. A consequentialist would find no reason to quarrel with him and it's hard to imagine a deontologist who would oppose the idea that the situation "has something to say" (Buber 1967, p. 617). He should also face little resistance with his view that one can never deal with a situation as one ought to; one is never "done" with a situation (Buber 1930, p. 163), i.e., it is humanly impossible to respond to the demand of a given situation in a way that would be (morally) impeccable. But as it will become clearer in the following, Buber's view of the situation is more radical and central to his moral philosophy.

The I-It and the I-You

Buber's opus magnum is generally considered to be his book *Ich und Du*, translated into English as "I and Thou".³² Its title is at the same time one of the central conceptual innovations of Buber and points to the singular position that he takes with regard to the subject-object dichotomy. In one sense he seems to suspend the

before nor is it composed of sounds that have ever been said. It can neither be interpreted nor translated, I can have it neither explained nor displayed; it is not a what at all, it is said to my very life; it is no experience that can be remembered independently of the situation, it remains the address of that moment and cannot be isolated, it remains the question of a questioner and will have its answer" (Buber 2002, p. 14) The translation of this passage seems problematic in some parts. The original German term „Sich-stellen“, which is here translated as „presenting ourselves“ has the connotation of confronting or facing up to something. When someone presents herself in this sense oft he word, she faces a challenge and takes responsibility. Similarly, the original German of the last sentence „it remains the question of a questioner and will have its answer“ is better translated as „it remains the question of a questioner and wants its answer“ (Buber 1930, p. 156). The original German passage thereby emphasises the urgency and demand that an agent is faced with in a given situation instead of depicting one, in which an answer is prepared and will be given eventually.

³² It is a matter of debate whether the German original of ‚Du‘ should be translated as “You” or “Thou”. The original translation of Ronald Gregor Smith used the rather archaic term “Thou” while the Walter Kaufman translation of 1975 preferred the term “You” (Kaufman 1996, p. 1). Although there is a religious element in Buber's understanding of “You” (he postulates an absolute, never-ending, eternal ‚ewiges Du‘, which can be interpreted as God) which makes the usage of “Thou” with its biblical connotations possible. But there is a directness and an intimacy about the everyday “Du”, which captures Buber's emphasis on the precedence of everyday situations better than “Thou”. Therefore I shall follow Kaufman in his translation of “Du” as “You”.

separateness of subject and object when he postulates that there is no I by itself (Buber 1970, p. 54) and that the I is necessarily always in a relation. The duality of subject and object is for Buber the foundation of philosophy, even if “the philosophical act should end in a vision of unity” (Buber 1953a, p. 526, translation B.S.). While in a religious experience the twofoldness of I and You grounds the nature of the encounter, philosophy divides the encounter into two essentially separate beings; one that experiences and perceives and one that can be nothing but the object of experience and perception (ibid.).

The mistake of the philosopher is to think that “he wishes to and can philosophise in his concrete situation” (Buber 1953a, p. 532). Buber regards the Cartesian method as an example of such an inadmissible abstraction. He dismisses the Cartesian Ego as “the product of a threefold abstracting reflexion”, since “ego cogito” can be understood as “I am the one who has consciousness” (ibid.). Buber interprets the Cartesian ego as the owner of consciousness and thereby of self. His critique echoes Hume’s critique of the idea of the self and rejects the notion that consciousness is part of what is experienced in the concrete situation. In the second step of the abstraction, it is postulated that there has to be a subject that is the owner of this consciousness; i.e., the ego and finally this ego is identified as the person or the self:

An ‘I think that’ is produced from the ‘that’ of the concrete situation, which encloses sensation and that, what is sensed, imagination and that, what is imagined, thinking and what is thought; i.e., a subject thinks this object, then the fundamentally indispensable ‘that’ (or something or it) is left out and now we acquire the assertion of this person about himself: therefore I (not the subject, but the living person who talks to us) real existence, as it is supposed to be the ego, where this existence is involved (Buber 1953a, p. 533).

In contrast to Descartes, Buber argues that the I cannot be experienced in such an abstraction and is without philosophical access. Buber’s I can only be experienced as

existing in a genuine relationship with a You (ibid.); i.e., fundamentally in a relation. Philosophy has the concrete situation as a starting point but runs the risk of losing the access to the concrete by way of philosophical abstraction (ibid.). This division of the self in perception and in action corresponds to a reflexive alienation that Buber interprets Henri Bergson as discussing; “In experience saturated by reflexion, the person sees himself acting, i.e., the duplicity of perceiver and perceived has found its way to the human person itself” (Buber Bergson, p. 1074).

Encounters and relationships are qualitatively different concepts in Buber’s terminology. Encounters are moments where the I is part of the couple I-You and they are characterised by what characterises the realm of the I-You: They are timeless or rather they are characterised by discontinuity, they are “actual” in both senses of the word (Buber 1963, 603). A relationship (*Beziehung*) on the other hand “opens up the possibility – only the possibility but it does open it – of latency.” (Buber 1963, p. 603, translation B.S.). Every I-You encounter is bound to return to an I-You relationship but such relationships may differ according to the potentiality they have of turning the other into a You.

I-It and I-You are the two ways of approaching being and as such they are the two modes of human existence; “The world is twofold for man in accordance with the two basic words he can speak” (Buber 1970, p. 53). The I of the human being is also twofold and is different depending on whether she is in the I-It mode or I-You mode of being (ibid.). It and You are part of what Buber calls basic words and they bring something into existence rather than denominating something that already exists (ibid.). That the I is twofold (*zweifältig*) should not be misunderstood to mean that it is fragmented but rather that it is relational and its momentary being depends on which relation it is in. A human being is part of either one or the other of the two

possible pairs, “[...] life is by its very nature an oscillation between You and it” (Buber 1970, p. 101).

Broadly, the realm of the I-It can be identified as that of experience while the realm of the I-You is that of relation (Buber 1970, p. 56). This can only be a broad approximation for the boundary between the I-You and the I-It is not a rigid one another pair of decisive criteria is presence and objectification:

For the real boundary, albeit one that floats and fluctuates, runs not between experience and non-experience, nor between the given and the not-given, nor between the world of being and the world of value, but across all the regions between You and It: between presence and object”³³ (Buber 1970, p. 63).

Whether an I is in an I-It or an I-You relationship with the other it faces does not depend on the other in a way that they ‘provide’ what is needed for the one or the other. Rather, “Every You in the world is doomed by its nature to become a thing or at least to enter into thinghood again and again” (Buber 1970, p. 69). Conversely, everything can become the You in an I-You relationship (ibid.). It is not the case that some objects, people, artworks or natural landscapes are more or less suitable for becoming either It or You. Furthermore, “it is not always as if these states took turns so neatly; often it is an intricately entangled series of events that is tortuously dual” (Buber 1970, p. 69). It is the case however, that a disproportionate weight on the I-It world will make it less likely that the individual will be able to encounter the other as a You; “The development of the function of experiencing and using comes about mostly through decrease of man’s power to enter into relation” (Buber 1958, p. 43).

³³ The original term for “presence” is *Gegenwart* and means not only presence as opposed to past or future but also the present being. The word is partly made up of *gegen*, meaning facing, vis-a-vis, opposite and thus the word includes three of the aspects, which are central for an I-Thou relationship; the timeless presence and being faced by the other. The original term for “object” is *Gegenstand*, While it also includes the word *gegen*, it is the *-wart*, which could be translated as “becoming” as opposed to the *-stand*, which is related to “standing” which gives *Gegenwart* the third connotation of Buber’s understanding of an I-Thou relation as a dynamic process, while *Gegenstand* is the fixed object which can be experienced, described and analysed. This play of words is unfortunately lost in translation.

A fundamental difference between the realms of the It and the You is that while the former can be perceived and described as an occurrence, which can be located in space and measured in time, the latter is not. Buber writes that the world of the It, as opposed to that of the You has coherence, continuity and relation in space and time: “The It-world hangs together in space and time. The You-world does not hang together in space and time” (Buber 1970, p. 84) and: “The world of It is set in the context of space and time. The world of Thou is not set in the context of either of these” (Buber 1958, p. 33).

Buber remarks in a number of passages that the I-It allows of degrees according to how near or how far from the I-You one is. The difference between the realm of the I-It and that of the I-You is a difference in kind,

[b]ut its [I-It relations, B.S.] highest stage is unmistakably set in contrast to the realm of the I-Thou relation, since even there an objectification prevails for which there is no room in this relation. A being to whom I really say ‘Thou’ is not for me in this moment my object, about whom I observe this and that or whom I put to this or that use, but my partner who stands over against me in his own right and existence and yet is related to me in his life (Buber 1964, p. 21).

There are degrees in the amount of objectification in an I-It relationship; a situation in which the I is part of an I-It couple can be nearer or further away from an I-You relationship (Buber 1964, p. 57). The highest degree of the I-It relationship, the degree of highest concentration is philosophical knowledge, since it extracts “the subject from the I of the directly lived togetherness of I and It” (Buber 1953a, p. 537) and transforms the It into the abstract, separate object existing by itself (ibid.). Nevertheless, objectification is not something that can ever be total, even if two people are bent on watching and analysing each other, “the decisive thing is not that one turns the other to be his object but rather the fact that he doesn’t succeed completely and why he doesn’t succeed completely” (Buber 1954, p. 275, translation

B.S.). The only reason that Buber gives for such a human immunity from complete objectification, is the “hidden action of my being” (Buber 1954, p. 275, translation B.S.). Despite the dangers involved in remaining in I-It relations to such a degree as to lose the access to the realm of the I-You, total objectification, in Buber’s view, is not possible.

The Realm of the I-It

The realm of the I-It is that of goal-directed attitudes, judgements and actions, of perception, emotion, imagination, volition, sensation and cognition (Buber 1970, p. 54). In an I-It relation, the I experiences the other as an It, i.e., it experiences, perceives and recognises aspects of the other as an object. Although critical of an excess of the element of I-It in everyday life, Buber emphasises that it should not be misunderstood as the evil or even as the less desirable of the two realms. The I-It only becomes problematical if it presumes to be the totality of being, similar to matter if it is considered to be the totality of being, as in materialism (Buber 1995, p. 49). Related is Buber’s criticism of reason, if it assumes to be the sole criterion of judgement:

[...] reason seems to me to take different attitudes in different times and circumstances. Either it knows itself as belonging as a part to the total being of the human person, and is active in full co-operation with the other properties and functions, and can in just this sense have a significant, yes even a leading, share in the intercourse of this person with other persons. Or it claims for itself the supremacy to which all the other faculties of man have to subordinate themselves. If it makes such a claim, then it appears to me presumptuous and dubious (Buber 1964, p. 53).

It is important to remark and it might help to distinguish the two realms that Buber would count the relations advocated under a care-ethicist perspective not, as might be

expected as I-You but as I-It relationships. Responding to Levinas, who affirms the importance of care, Buber remains critical towards care as an access to the other:

The reality of experience seems to me that he who has this access already, will find it also in the care that he practices, he who does not have it already, may clothe naked people and feed hungry people all day, it will [nevertheless] remain difficult for him, to speak a genuine You. If everyone was well-dressed and well-nourished, the real ethical problem would only then appear properly clearly (Buber 1963, p. 620, translation B.S.).

Heidegger's understanding of *Fürsorge* (care), which seems to provide an access to the other is also dismissed by Buber "[f]or the relation of solicitude which is all he considers cannot as such be an essential relation, since it does not set a man's life in direct relation with the life of another, but only one man's solicitous help in relation with another man's lack and need of it" (Buber 1938, p. 201). Care and solicitude cannot be the origin of the approach to the other but they can arise as a by-product, so to speak, thus making the relationship primary. Buber's discussion of the relation that involves care is an example of his ontologically radically relational position and helps to explain the difference between an approach involving care and assistance, in which the subject-object relationship characteristic of the I-It sphere remains intact:

In mere solicitude man remains essentially with himself, even if he is moved with extreme pity; in action and help he inclines towards the other, but the barriers of his own being are not thereby breached; he makes his assistance, not his self, accessible to the other; nor does he expect any real mutuality, in fact he probably shuns it; he 'is concerned with the other', but he is not anxious for the other to be concerned with him. In an essential relation, on the other hand, the barriers of individual being are in fact breached and a new phenomenon appears which can appear only in this way: one life open to another – not steadily, but so to speak attaining its extreme reality only from point to point, yet also able to acquire a form in the continuity of life; the other becomes present not merely in the imagination or feeling, but in the depths of one's substance, so that one experiences the mystery of the other being in the mystery of one's own. The two participate in one another's lives in very fact, not psychically, but ontically" (Buber 1938, 201). Care remains in the I-It sphere, the other, when approached with care, is not a You (Buber 1938, 204).

The realm of the I-You

It is easy to misunderstand the realm of the *You* as something esoteric, spiritual or fuzzy. One reason for this is methodological, since Buber insists that the encounter between I and You is beyond time and impossible to convey with words (Krone 1993, p. 47). Buber appeals to the experiences of his partner in dialogue rather than attempting an argument to prove the existence of the I-You relation or to demonstrate its characteristics. This remains a methodological problem both in trying to understand his position as well as in evaluating it.

Another reason why the *You* can appear as esoteric is Buber's negative description of it as having another basis than the realm of the *It*, which is characterised by perception, sensation, emotion, cognition, judgement, volition and action (Buber 1970, p. 54). As this list seems pretty comprehensive, what might remain outside its scope seems to be outside what can be accessed by scientific or philosophical methods. But the point that Buber is trying to make is that the mode of being in which a human being is when in an I-It relation is that of experiencing an object as an object, which borders on other objects; "It is only by virtue of bordering on others" (Buber 1970, p. 55). His point is related to Kant's categories of intuition, with which the human mind draws borders and thereby orders the world: "Those who experience do not participate in the world. For the experience is 'in them' and not between them and the world" (Buber 1970, p. 56).

What can happen as an encounter between the I and the You can take place in three spheres – in life with nature, in life with human beings and in life with "spiritual beings"³⁴ (*geistige Wesenheiten*). (Buber 1970, p. 56). What is meant with

³⁴ Once again the English translation is problematical. Both Kaufman and Smith translate *geistige Wesenheiten* as „spiritual beings“. The word „*Wesenheit*“ involves the elusive *Wesen*, which is central

the latter area is not that of spirits and esoteric phantasma but the area of human creativity; most paradigmatically art.

The encounter between I-You is not epistemological; it does not serve to provide information about the other. What could be regarded as objective knowledge or what can be thought of as contributing to it belongs to the realm of the I-It. Buber is critical towards the high pedestal that rationality or reflection has been given traditionally but he is just as critical towards intuitionism and denies that intuition is an epistemologically promising route (Buber 1963, p. 601). Nevertheless, the switch from the I-It to the I-You, as Buber describes with the example of watching a tree does not involve a loss of the mode of experiencing the tree as an It; “Rather is everything, picture and movement, species and instance, law and number included and inseparably fused” (Buber 1970, p. 58).

The realm of the I-You is the realm of encounter and of relations, not of emotions or of the “unconscious”.³⁵ The so-called “inner life” of consciousness and emotions can also be encountered as an It; as a personal play set up for one’s amusement; “here one enjoys one’s inclination and one’s hatred, pleasure and, if it is not too bad, pain” (Buber 1970, p. 93). Shunning the so-called “outer life” of institutions and turning to one’s personal, mental life to escape the alienation that is so typical of modernity, is just to exchange one type of alienation with another;

to Buber’s ontology in general and his philosophical anthropology in particular and both translators use the same term (being) for both *Wesen* and *Wesenheit*. One can only speculate why Buber chose to use add *-heit* to *Wesen*, a suffix, which corresponds more or less to the English *-ness*. It is probably meant to stress the abstractness of *geistig* beings. More problematical is the translation of *geistig* as “spiritual” however for although it is common to translate *Geist* as spirit, as in translations of Hegel, the word *geistig* has a different, almost opposite connotation to that of spiritual. It denotes rather the cognitive aspects of the mind, as exemplified in the German term for the philosophy of mind; *Philosophie des Geistes* or *Geisteswissenschaften*, which corresponds to Humanities without Arts.

³⁵ “But the interhuman I-Thou relation does not belong to the unconscious, even in its most exclusive form, although its roots, of course, are sunk in the ‘unconscious’, that is, in the ground of being of the person. The consciousness of the I-Thou relation is a highly intensive one; but it is a direct, an elementary consciousness. It does not make itself an object; it does not detach itself from itself; its knowing about itself is given with its being” (Buber 1964, p. 39).

“Institutions yield no public life; feelings, no personal life [...] And once one has learnt, like modern man, to become greatly preoccupied with one’s own feelings, even despair over their unreality will not easily open one’s eyes; after all, such despair is also a feeling and quite interesting” (Buber 1970, p. 94).

“Buber’s ‘I-You’ relation is one that can only be of short duration, but its significance is that after one has had an ‘I-You’ relation with the divine, the ‘It-World’ is transformed. There are, so to speak, two sorts of ‘I-It’ relations: mere ‘I-It’ relations and transformed ‘I-It’ relations” (Putnam 2008, p. 63). The I remains real; it is still in a relation as it is always in an I-It relation if it is not in an I-You relation, but the categorical difference involves a consciousness of the I and of separation:

But the I that steps out of the relational event into separation and consciousness of separation, does not lose its reality. In other words, as it is said of the supreme relation and may be used of all, ‘the seed remains in it.’ This is the province of subjectivity in which the I is aware with a single-awareness of its solidarity of connexion and its separation (Buber 1995, 66 / Buber 2004, 52).

“Experiencing the other side” (Buber 1926, p. 114) has a similar transformative effect on the everyday life in which the other is experienced as an I. Buber gives two examples, one of domination and the other of (erotic) love, in which the actor experiences moments of diffusion, in which he perceives his own actions both as agent and as receptor (ibid.). Against a possible misunderstanding Buber emphasises that the agent who has had this experience will not continue to experience all of his actions in such a dual perspective. Nevertheless, as in the case of the transformed approach to the It, which has once been a You, “the one extreme experience makes the other person present to him for all time. A transfusion has taken place after which a mere elaboration of subjectivity is never again possible or tolerable to him” (Buber 1926, p. 114).

Buber emphasises that this phenomenon, which he calls ‘inclusion’ (*Umfassung*)³⁶ should not to be confused with empathy. Empathy involves “the exclusion of one’s own concreteness” (Buber 1926, p. 115) as it is the simulation of what it feels like to be the other. Buber’s inclusion (*Umfassung*) is not the substitution of one first-person-perspective with another but rather an extension. The elements of (*Umfassung*) are

first, a relation, of no matter what kind, between two persons, second, an event experienced by them in common, in which at least one of them actively participates, and, third, the fact that this one person, without forfeiting anything of the felt reality of his activity, at the same time lives through the common event from the standpoint of the other (Buber 1926, 115).

Dialogue, I-You relations and affirmation are central concepts of Buber’s thoughts. They are closely related but distinct concepts and should be properly distinguished. Affirmation is not a prescription but the natural effect of being in a dialogue. There are two ways or rather two degrees of approaching the other affirmatively and these are comparable with the more common concept of empathy without being identical with it. *Vergegenwärtigung*, (literally: to make present) is based on a skill which everyone has in some measure and which Buber calls *Realphantasie*. Buber defines this skill as “the ability to bring to mind a reality, which exists in a given moment but cannot be experienced with the senses” (Buber 1950, p. 422, translation B.S.). *Realphantasie* is similar to but not identical with empathy (*Einfühlung*). Buber understands empathy as experiencing the consciousness of the other, while losing a sense of one’s own self (Buber 1926).

The difference between *Realphantasie* and empathy appears also in Buber’s discussion of Bergson’s concept of intuition and points to the “primal problem of the contradiction between being and perception/knowledge” (Buber 1962, p. 1074),

³⁶ “Inclusion” is perhaps not the best translation for *Umfassung*, which does include the element of including something else in the realm in question but extends the realm rather than incorporating the other: “Enclosure” might be more suitable to involve this connotation.

which is the basis for both the need and the possibility for empathy. Empathy may diminish the gulf between my perception of the other and her being but does not annul it (ibid.). Buber regards the tension between the perception and the being as opening up the possibility of turning into an I-You relation, in which the You, “which I encounter is no longer a sum of perceptions, no object of *Erkenntnis* but a substance that is experienced in giving and taking” (Buber 1962, p. 1074).

Buber differentiates between the other and the encounter and it is the latter, not the former that is foundational for the I: “Das Ich wird am Du”, “The I becomes by/with the You” – this sentence, Buber complains, has been misunderstood to mean that the other in the encounter is what makes the I to what it is. The other is indeed central, it is the other who makes an encounter possible and it is the other who co-determines the nature of the encounter. Nevertheless, it is not the other herself that is the foundation for the I of this particular encounter but the fact that the I is in an attitude of I-You with the other: “Only in the relation is he my You, outside the relation between us this You does not exist. [...] Neither is my You identical with the I of the other nor is his You identical with my I. What I owe the person of the other is that I have this You; but my I – which should be understood as the I of this I-You relation – I owe to the saying of you, not to the person to whom I say You” (Buber 1963, p. 596).

Self and Moral Agency in Buber

The position and nature of the (human) individual is the binding focal point of Buber’s thought. As is the case with many of his other central concepts however, his postulations about personhood, self and agency are complicated by the variety of

aspects that Buber considers, as well as the variety of terms he uses to differentiate these aspects. Terms that appear around the theme are “*Persönlichkeit*” (personality), “*Charakter*” (character), “*Person*” (person) and “*Subjekt*” (subject) (Krone 1993, p. 137) as well as “*Eigenwesen*”. Personality denotes the unique living figure, which is both mental and physical (Buber 1939a, p. 123) while character denotes “the link between what this individual is and the sequence of his actions and attitudes” (ibid.). It is only character that can be the task of the educator to address while she has no influence on the development of personality (ibid.).

Despite his adoption of such terms as personality, character, person or subject, which are traditionally closer to the idea of the transcendental ego, Buber argues that the phenomenon of continuity in self-consciousness is not a sufficient reason to postulate an I, which exists without being in a relation:

Between the I that in a given moment detaches itself from the other existing being and the I that in another given moment turns to the other existing being, there exists, incontestably, a special kind of continuity that is preserved despite all discontinuities; and it is this which one customarily designates as self-consciousness. But I do not see that this fact justifies the acceptance of an isolated I that stands over against neither a Thou nor an It and is not even comprehended in the transition from the one to the other relationship to being (Buber 1964, p. 28).

The view that the nature of the I is mistaken if it is regarded as something akin to a transcendental ego, i.e., as the unified synthesiser and subject of experiences and action rather than as a relational entity is not new to Buber: Michael Bakhtin, who applied the term of dialogue to literary theory writes: “A single person, remaining alone with himself, cannot make ends meet even in the deepest and most intimate spheres of his own spiritual life, he cannot manage without another consciousness. One person can never find complete fullness in himself alone” (Mikhail Bakhtin, quoted in Friedman 2005, p. 31). The reason for this incapability is that personality

means neither Descartes' solipsistic I nor an object but another subject: "The depiction of personality requires addressivity to a thou" (ibid.). Buber himself quotes Ludwig Feuerbach as advocating a position close to his own:

The individual human being by himself does not have the essence (*Wesen*) of the human being in himself, neither in himself as a moral being (*Wesen*) nor in himself as a thinking being. The essence (*Wesen*) of the human being exists only in community, in the unity of the human with the human – a unity, which bases itself only on the reality of the difference of I and You (quoted in Buber 1948, 342; translation B.S.).

What distinguishes Buber from such comparable positions is not only his point that the I does not exist by itself but only in relation to something or someone else.

Building on that premise, he additionally postulates that there are two distinct ways of being in a relationship: "There is no I as such [*Ich an sich*] but only the I of the basic word I-You and the I of the basic word I-It" (Kaufmann 1970, p. 54). The realm of the It; i.e., the relations in which the I is the I of the word pair I-It is qualitatively different than that of the I-Thou or I-You. According to the realm, in which an agent is her perception of the world as well as her being will be different, so that Buber's epistemology and ethics are interconnected; it's the (moral) stance of the agent that determines her modus of perception and her idea of truth (Reichert 1996, p. 42). There are three such stances, two of which correspond to the I-It world and consider the other as an object to be studied and perceived. The distinguishing trait between an I-It moment and an I-You moment is experience. Buber argues that the conscious mental life is largely but not exclusively composed of inner and outer experiences; i.e., of the conscious perception of things and events 'outside' – in the world or 'inside' in one's own stream of consciousness. This sort of conventional experience, which in Buber's terminology are the experiences of *It* is not the only type of event. The way in which an I and a You face each other is of another kind, it

is not an experience of something else but it is an encounter, an event, in which one partakes (Buber 1922, quoted in Reichert 1996, p. 45).

Whether the I-You relation can take place is a matter of “grace and will” (Buber 1995, p. 77) or even simply of grace (Buber 1979, p. 62). To emphasise that what he means by “will” is not the product of the cognitive capacities of highly intellectual or spiritual persons, he also uses the term “heart-will” in another passage (Buber 1964, p. 36). In another passage he changes the terminology “Meetings stand – as I have repeatedly indicated – under freedom and under grace, therefore not under an ‘unbending law’” (Buber 1964, p. 20). The term “grace” can be understood as referring to those elements of a given situation that will or will not allow it to turn into an encounter if these are beyond the individual’s influence. Buber therefore remarks that “We have to be concerned, to be troubled, not about the other side but about our own side, not about grace but about will. Grace concerns us in so far as we go out to it and persist in its presence; but it is not our object” (Buber 1995, p. 77 / Buber 2004, p. 62). This should not be misunderstood to mean that the attitude of the agent is under her own control however. Whether an agent is capable of demonstrating the openness and attentiveness necessary for an encounter in the genuine sense is not simply a willed decision but also grace.

Personhood and *Eigenwesen*

Buber’s concern is predominantly with human beings and his understanding of personhood aims at understanding persons with flesh and blood. When considering a person his tendency is to understand her as an agent rather than as the subject of perception or knowledge. The anthropological view of man in Buber’s thought is not

the subject of epistemology but rather the concrete human being (Buber 1936, p. 209). This is the case also when the individual is considered as a moral agent; the *Umkehr*, a reversal that can be interpreted as the begin or re-begin of moral action is “as little a ‘mental’ event, as birth or death of man; it happens on the whole person and it doesn’t happen in the relation of the individual with himself but in the simple reality of primary reciprocity” (quoted in Reichert 1996, p. 212, translation B.S.)

The two realms of being, the realm of the It and the realm of the You correspond to two modes of being of the I. The I that is in a relation with an It is not the same I as that in a relation with a You (Buber 1995, p. 65 / Buber 2004, p. 51). The I in a relation with an It is an ‘*Eigenwesen*’,³⁷ while the I in a relation with a You is a ‘Person’. There is a specific self-consciousness proper to the *Eigenwesen* as the subject of experience and action; i.e., as an agent while the self-consciousness of the person is “subjectivity” in a more abstract sense (ibid.). The notion of self-consciousness of the *Eigenwesen* is familiar; it is the conception that an individual has of herself as a particular being with particular qualities; of its quiddity (“*Sosein*”).³⁸ What the person conceives is not the qualities that set her apart from others and thereby determine her individuality as something separate and unique. It is rather the consciousness of her taking part in being but unlike the case in the Buddhist tradition this does not involve a loss of self. On the contrary, the person retains a consciousness of both her being distinct than others and of being a particular being but she does not distance herself from Being – with a capital B – as the *Eigenwesen* does.

³⁷ Buber calls the I of the pair I-It “*EigenWesen*”, which denotes something like “being by/for itself”. It has been translated into English as “individuality” (Buber 2004, p. 51) but this translation seems to be unsatisfactory. The entity in question is not a state of being or an attribute so a grammatically more appropriate translation would be “individual”, which carries with it too many connotations to be used in place of the Buberian “*Eigenwesen*”. I shall therefore use the German original.

³⁸ Buber expands the term ‘*Sosein*’ with “So-und-nicht-anders-seienden [*Wesen*]” (Buber 1995, p. 66); being in a particular way and in no other way.

The modus of being that is the *Eigenwesen* is part of Buber's critique of modernity. The *Eigenwesen* has succumbed to the idea of individuality gone astray by creating an image of herself as a particular and separate creature with a number of concrete and individualising properties. Buber is severe in his criticism of the *Eigenwesen*'s tendency for self-deception: "For to know itself means basically for it (for the most part of it) to establish an authoritative apparent self capable of deceiving it"³⁹ ever more and more fundamentally, and to procure for itself, in looking to and honouring this apparent self, the semblance of knowledge of its own being as it really is" (Buber 2002, p. 52). The *Eigenwesen* is more concerned with what is part of her wider net of affiliation or ownership; her species, race, productions or genius (Buber 1995, p. 67).

This makes clear that what Buber is hinting at with his understanding of personhood is not part of the discourse around communitarianism. On the contrary, Buber is opposed to one type of collectivism as he is to individualism⁴⁰. Buber's critique of Karl Marx is that he did not incorporate the "element of the real relation between the real different I and You [as Feuerbach had done, B.S.] and precisely because of this did he oppose unrealistic individualism with a just as unrealistic collectivism" (Buber 1948, p. 342). Both individualism and collectivism are abstractions, which misconceive the *Wesen* of human beings as fundamentally and ontologically relational.

Buber's particularism aims to provide an opportunity to avoid the pitfalls of both atomistic individualism as well as the diffusion of collectivism. Both

³⁹ A correct translation of the original text would have been „itself“; the *EigenWesen* produces an apparent self with which it deceives itself.

⁴⁰ "I do not consider the individual to be either the starting point or the goal of the human world" (Buber 1930b, p. 82)

individualism and collectivism are contemporary reactions of human beings to the particular crisis of modernity, argues Buber; “[...] if individualism understands only a part of man, collectivism understands man only as a part: neither advances to the wholeness of man, to man as a whole. Individualism sees man only in relation to himself, but collectivism does not see man at all, it only sees ‘society’” (Buber 1948, p. 237). The collective of people, which could be regarded as a “we” for Buber is characterised by the same ontological direct connection as two individuals in an I-You relation are. Only such people are capable of speaking “we”, Buber argues, while other collectives are simply “one” (Buber 1938, p. 373).

The particular danger of being a part of a community is the latter’s provision of security and pre-established patterns of perception, evaluation and behaviour.⁴¹ Buber believes that the collective element preserves the individual from loneliness, a sense of being lost and of *Weltangst* and while this function of the collective is not to be rejected as such, it may lead to fewer encounters and “life between person and person” (Buber 1954, p. 272).

Buber’s distinction of person versus *Eigenwesen* does not propose distinct categories of being but rather two poles of humanity between which an individual might be situated (Buber 1995, p. 67); “No man is pure person and no man pure individuality. None is wholly real and, and none is wholly unreal. Every man lives in the twofold I” (Buber 2004, p. 53). The categories of Person and *Eigenwesen*, which correspond to the You and the It thus also correspond to what Buber terms the real and the unreal. The tendency of the *Eigenwesen* to establish for herself a world of separate objects, in a given order and with a given meaning is, according to Buber,

⁴¹ Pierre Bourdieu provided an empirical account and sociological analysis of correlations between values and social origin and demonstrated that the latter is one of the main factors to explain taste and distinction (Bourdieu 1982).

an estrangement from reality which is accessible to the person. The world of the *Eigenperson* is accordingly necessarily one of self-deception. The more this tendency of the *Eigenperson* takes over in an individual or in humanity, the more unreal the existence of the I becomes (Buber 1995, p. 68). This “reality” should not be understood as something existing independently of the I however; Buber is not endorsing a representational theory of mind, in which the perception of the outer world may be more or less congruent with its possible mind-independent qualities. Rather, Buber seems to hold that there are grades according to which one can be more or less successful in answering the demand of a given situation and it is the success in this that makes a situation more or less real. The individual thus partakes in the coming to being of reality (Krone 1993, p. 77).

Buber’s dictum that the I does not exist by itself but is rather always in a relation finds its negative formulation in his objection to what Friedman terms psychologism: “What is essential is not what goes on within the minds of the partners in a relationship but what happens between them. For this reason, Buber is unalterably opposed to that psychologism which wishes to remove the reality of relationship into the separate psyches of the participants” (Friedman 2005, p. 29). While discussing Max Scheler’s anthropology, Buber explicitly rejects the view that the nature (*Wesen*) of human beings could be grasped by investigating the mental life of the individual or his self-consciousness. The crucial characteristic in Buber’s anthropology is rather the specific nature of human beings’ relations to things and creatures.⁴² Buber rejects both materialism and idealism;⁴³ reality is to be found in

⁴² “Und sowohl hier wie dort ist das *Wesen* des Menschen nicht von dem aus zu erfassen, was sich im Innern des Einzelnen abspielt, und nicht von seinem Selbstbewußtsein aus, das Scheler für den entscheidenden Unterschied zwischen Mensch und Tier hält, sondern von der Eigenart seiner Beziehungen zu den Dingen und den *Wesen*“ (Buber 1948, p. 401).

⁴³ Buber does not use these terms explicitly but draws the picture of a human being torn between two

neither one or the other and the belief that it should be so leads to alienation from “real life”, which Buber locates in the In-between.

The ‘In-between’ thereby becomes one of Buber’s central terms. It denotes the space, in which existence and events happen and is the ontological centre of Buber’s thought. Unfortunately it is also something of a black hole, since – to my knowledge – Buber does not provide a separate discussion of the ontological nature of the In-between and only remarks to it in passing. What takes place “between man and man”⁴⁴ is distinctly not psychological: “One has to beware of interpreting emotional motifs into such fleeting and nevertheless consistent events [examples of genuine encounters, B.S.]; what appears here is not accessible to psychological concepts, it is something ontic” (Buber 1948, p. 406). Studying an encounter, both the physical and the psychological elements can be extracted; however there still remains the encounter, the dialogue itself, for the localisation and understanding of which Buber deems the category of the ‘In-between’ as necessary (Reichert 1996, p. 148). Buber’s remarks on the subject are thus negative and not positive; we are told as what not to understand it but it’s not clear what ontological status it should be accorded. The reason for this omission however is arguably not a deliberate neglect but a lack of perceived importance.

For Emmanuel Levinas the ontology of the In-between is not only no challenge, it is characteristic for the ontological approach of the day. Interpreting Buber, Levinas argues that “the ontological domain is not a block of being but an event” (Levinas 1963, p. 124). It should not be understood as a space independent of

ways of interpreting reality; once in purely physicalist terms in which “the I is contained in the world and that there really is no I” (Buber 1970, p. 121) and once as a product of consciousness, where “the world is contained in the I, and that there really is no world” (ibid). As long as the alienated human being can believe in one of these, she can interpret the world so as to refrain from recognising his alienation but not if she considers both perspectives at once (Buber 1970, p. 122).

⁴⁴ This is the term used in the English translation of Buber’s term “Zwischenmenschlich”, which is part of the title of one of his central texts, “Elemente des Zwischenmenschlichen” (Buber 1954 and Buber 2002)

I and You, where they might meet but rather the other way around; it's the encounter of I and You that produces the In-between, ever anew in every fresh encounter (Levinas 1963, p. 124). The reason why this approach to being is characteristic for the age is that it is involved in the dismissal of Being as "the content of being or as a material realisation or a 'narrative' being", which, for Levinas, is the distinguishing mark of the ontology of the time (Levinas 1963, p. 125). He thus interprets Buber's In-between as filling a gap in a new approach to ontology.

To Levinas' interpretation of the In-between as "the concept of the foundation and ultimate structure of being" (quoted in Buber 1964, p. 23), Buber replies negatively and practically, claiming that "we cannot do without this category for a full comprehension and presentation of what passes between two men when they stand in dialogue with each other" (Buber 1964, p. 27). What he means by the category of the In-between becomes a little clearer in a passage in which he tries to respond to the criticism that the concept has received:

I take my starting-point from a simple real situation: two people are taking part in a real conversation. I want to map the facts of this case. It turns out that the usual categories are not sufficient for this. I observe: firstly the 'physical' phenomena of the two speaking and gesticulating people, secondly the 'mental' phenomena of what takes place 'inside them'; but the conversation itself, which carries meaning, in which the acoustic and optical occurrences add themselves, which goes out of the souls and which is mirrored in the souls, has remained unrecorded. What is its nature, what is its location? My survey cannot do without the category, which I term the In-between (Buber 1963, 605, translation B.S.)

The Genesis of the I-It and the I-You

Although it would seem that the capability of an individual to immerse herself in an I-You encounter must be the result of experience, wisdom, choice, and courage and thereby to be acquired only gradually and late in life, Buber's understanding of

encounter takes the opposite route. The genetically primary state of being is not detached but relational. In a rather questionable passage, Buber describes this development as a quasi phylogenetical one:

It may be supposed that characterisations and ideas, but also representations of persons and things, have been taken out from representations of incidents and situations that are specifically relational. The elementary impressions and emotional stirrings that waken the spirit of the 'natural man' proceed from incidents – experience of a being confronting him – and from situations – life with a being confronting him – that are relational in character (Buber 2004, p. 22).

The reason that this passage is questionable is not the idea that the genetically primary state is the relational one, while the separation of subject and object is secondary but rather the idea that this process is not only ontogenetical but also phylogenetical. As such, it is part of Buber's critique of civilisation in general and modernism in particular. If there are any anthropological grounds to suppose that the 'natural man' developed into the 'civilised man' in this way, Buber does not mention them. Therefore this Rousseauian idea can be dismissed as a romanticisation of the difference between the so-called 'natural man' and the 'civilised man'. Buber is aware and explicitly comments on about the speculative character of this idea however and turns to the study of the child (Buber 1995, p. 28 / Buber 2004, p. 26).

The child, according to Buber is born with the *You* as the default mode of being and of perception; it is basic and innate. To emphasise the importance of the primary *You*, Buber even adapts the biblical slogan "In the beginning is relation" (Buber 1970, p. 69). The capacity of being able to abstract an *I* from this relation and to come to consider the other as an object existing independently of this *I*, as one that can be perceived, handled and studied is ontogenetically secondary:

It is simply not the case that the child first perceives an object, then, as it were, puts himself in relation with it. But the effort to establish relation comes first – the hand of the child arched out so that what is over against him can nestle under it; second is the actual relation, a saying of *Thou* without words, in the

state preceding the word-form; the thing, like the I, is produced late, arising after the original experiences have been split asunder and the connected partners separated” (Buber 1995, p. 31 / Buber 2004, p. 25).

It would be to misunderstand Buber to interpret these passages as a yearning for an original and better state of affairs. The child is bound to the world in a way that the adult cannot regain but this does not mean that the child is in the same sort of relationship or encounter as the adult in the I-You state. Buber distinguishes “*naturhafte Verbundenheit*” from “*Beziehung*”, i.e: “natural connexion” (Buber 2004, p. 27) from “relation“, which is „*geisthaft*“ („spiritual“) ⁴⁵ (ibid.). The process, by which the primary relational state of the child is maintained although a sense of objects as separate from the I has additionally been acquired, provides the adult with her I, capable of being both in I-It and in I-You relations; “Man becomes an I through a You” (Buber 1970, p. 80), it is not a given.

The *Wesen* of a human being

One of the most challenging concepts in Buber’s thought is the *Wesen* of a human being. In Walter Kaufman’s translation of I and You, it is translated as “being” (Buber 1970). The word has many connotations however, which are not caught by this generic term. *Wesen* can be understood as an entity but also as the nature of an entity, for example the character of a human being. It can be understood as the essence of a thing; i.e., the quality or qualities that an entity must necessarily have in order to be that particular entity. Buber himself defines *Wesen* as “that, with which

⁴⁵ The translation of “*geisthaft*” as “spiritual” does not wholly cover the meaning. While the term “spiritual” has acquired an esoteric connotation and implies a realm of being, perception and communication that is emotional or goes beyond the physical, “*Geist*” and its variations in German denote a more cognitive or intellectual realm as opposed to a sensual or emotional one. If the child is to achieve a conversion of the original, natural bond with the world into a “*geisthaft*” relationship, then the process is one of exploration and abstraction (Buber 1995, p. 29 / Buber 2004, p. 27).

the person is singularly invested with, that, which he is destined to become.

Consciousness, with its plans and ponderings attends to it only occasionally, the subconscious with its wishes and contradictions does almost never so” (Buber 1957, p. 484, translation B.S.)

Replying to a critic who questions the coherence of the term, especially in the constellation “whole *Wesen*”, Buber attempts a description:

I enter into the act or event which is in question with all the available forces of my soul without conflict, without even latent much less perceptible conflict. [...] Note well, the resistance must certainly not be presupposed in any given situation; there are souls that have long since overcome analogous resistances and now are already capable of meeting as a whole the situation that accosts them; indeed there are souls of whom we don't not know that the battle within them has ever been fought through, yet whose wholeness nonetheless in an unforeseen situation begins forthwith to shine like the sun (Buber 1964, p. 52).

The term also appears in two central lines of Buber's thought; in his ontology of the I and his moral theory. It is the differentiating characteristic between the I in an I-It relation and the I in an I-You relation: “The basic word I-It can only be spoken with one's whole being [*Wesen*]. The basic word I-It can never be spoken with one's whole being” (Buber 1970, p. 54). The term receives a teleological connotation in that it points to a state of being, which is “whole”, where the different components of the self seem to be in harmony, not as “partial actions” (Buber 1970, p. 62) and turned to a common direction, to the Thou: “[...] the wholeness of the soul: I can only – to repeat ever again the same thing – that we can speak the true Thou only with the whole soul, where the stubborn contradiction no longer lurks in the corners” (Buber 1964, p. 82).

Buber's use of such a metaphysically rigid concept is challenging precisely because of his insistence that the I cannot exist on its own but only in a relation. Since the I is a different I in every relation and since that, with which it is in a

relation is changeable as well, it seems hard to imagine how there could be a *Wesen* involved. To make things even more difficult, Buber goes beyond postulating a *Wesen* and additionally supports a quasi-normative stance that seems to imply that there is a core of the *Wesen* that demonstrates *Stetigkeit*; continuity or steadiness (Buber 1954, p. 280). This stance is quasi-normative because Buber repeatedly remarks that the highest goal of every *Wesen* is to acquire “the highest quality, that of inner unity”⁴⁶ (quoted in Reichert 1996, p. 273). The main reason why unity is accorded such high value is Buber’s conviction that morality is more possible the higher the degree of unity and direction of the moral agent is. Thus what the immoral character and the amoral character need to acquire in order to turn into moral agents is unity (ibid.). Unity is the condition for moral action, “a decision that is made by just one part of the person is not a real decision” (quoted in Reichert 1996, p. 174). Another reason is Buber’s conviction that there is a special, unique and singular part that every being can and ought to play in creation. Every person born into this world represents something new, something that never existed before, something original and unique.

“Every man’s foremost task is the actualisation of his unique, unprecedented and never-recurring potentialities, and not the repetition of something that another, and be it even the greatest, has already achieved” (Buber 1996). Expressed in a more secular fashion, this stance supports the value of individuality against human tendencies towards laziness, conformity, fear, close-mindedness or inattentiveness. If we interpret Buber’s description of the Chassidic dictum that there is no universal way of the good or right life but that every Chassid must find her own way of life as an appeal to the moral agent, then there are some guidelines that can be postulated

⁴⁶ Once again this passage is part of Buber’s exposition of Chassidism but the great number of his own remarks in this regard seem to justify the view that he shares the Chassidic account.

despite Buber's repeated comments to the contrary. These can be understood as procedural criteria for an agent to find the individual mode of action by attaining a unity of being. The first of such guidelines is "Beginning by oneself"; i.e., to overcome inner conflicts, which are thought to be the basis of interpersonal conflicts: "Not to be preoccupied with oneself" is the principle that points to the ultimate goal of self-unification, which is not oneself but others and the world: "To begin with oneself, but not to end with oneself; to start from oneself, but not to aim at oneself; to comprehend oneself, but not to be preoccupied with oneself." (Buber 1996). Buber's focus is not on the good life as an aesthetic of existence but on the life that will contribute as much as possible to the improvement of the rest of the world. The last recommendation is one that invites the agent to humility and to locality. The agent should not divide the world into sacred and profane but approach everything with attentiveness and openness in order to be able to give them their due:

The people we live with or meet with, the animals that help us with our farm work, the soil we till, the materials we shape, the tools we use, they all contain a mysterious spiritual substance which depends on us for helping it toward its pure form, its perfection. If we neglect this spiritual substance sent across our path, if we think only in terms of momentary purposes, without developing a genuine relationship to the beings and things in whose life we ought to take part, as they in ours, then we shall ourselves be debarred from true, fulfilled existence. It is my conviction that this doctrine is essentially true (Buber 1996)

For our purposes we can take Buber's description of the trajectory of the human being along the Chassidic way of life as one way to determine which characteristics the successful moral agent ought to have.⁴⁷ It is something of a challenge to correctly interpret Buber's stance towards such characteristics. A virtue-centred approach could interpret them as virtues in disguise. One of these characteristics is

⁴⁷ It is difficult to explicate Buber's views without either using terms that he would object to or succumbing to a paraphrase, which does not go beyond a reformulation. In this example, Buber would object to the use of "ought", since he resists the idea of universalisability of individual ways of life. He might also be uncomfortable with the adjective "successful", which I use to delineate the area, in which a human being has accomplished being as authentic, attentive and open-minded towards the situation she finds herself in as possible.

Selbstbesinnung; it is one of the terms that seem to be positioned in the middle of the tension between Buber's relational ontology and the pleasure he takes in the idea of a constant and steadfast core of being. *Selbstbesinnung* could be translated as self-reflection, it involves a turning-to-oneself and is therefore a way of introspection. In order to fulfil the demand of responsibility towards the other in an encounter, the agent must have self-awareness about her tendencies towards missing the (morally) relevant factors of the situation (Krone 1993, p. 127).

The concept of responsibility is not a back door through which principles or calculations could be integrated into Buber's moral theory because it is closely connected with the concept of the situation and the responsibility one has is not towards an ideal or a principle but towards a very concrete speaker. Buber follows Max Stirner in his attack of

the substitute for a reality that is no longer believed: the fictitious responsibility in face of reason, of an idea, a nature, an institution, of all manner of illustrious ghosts, all that in its essence is not a person and hence cannot really, like father and mother, prince and master, husband and friend, like God, make you answerable (Buber 1936, 207 / Buber 2002, 53).

"Cold morality" as cool reflection on principles or maxims and their applicability is not possible or rather is neither reality nor morality. Responsibility without a receiver is an illusion (Buber 1936, p. 208 / Buber 2002, p. 53). In another passage, Buber dismisses the notion of "responsibility towards an idea" as "fictional, because the idea cannot call me to account, it cannot decide whether my responsibility exists by rights or not" (Buber 1963, p. 596). Similar to Georg Simmel's postulation of a necessary link between rights and duties,⁴⁸ Buber's concept of responsibility can

⁴⁸ Simmel's article on the sociology of poverty begins with a dictum that every duty of a human being corresponds to the right of another. His further explication, although sociological, could be mistaken for a sentence by Buber: "Vielleicht ist es sogar die tiefere Auffassung, dass es von vornherein nur Rechte gibt, dass jedes Individuum Forderungen - allgemein menschlicher und aus seiner besonderen Lage hervorgehender Art besitzt, die erst als solche zu Pflichten anderer werden." (Simmel 1906, p.1)

only be thought of as existing between someone who demands and someone of whom an answer is expected.

While singling out elements like attentiveness, open-mindedness etc. as the characteristics of a moral agent most suitable for a moral situation, he is critical towards the traditionally more popular reason:

That it is 'reason' that opposes me in evil that I do seems to me an inadmissible simplification. When I think about doing an injury to my neighbour who has vexed me, and I succeed in sensing somewhere in a corner of my being the injury that I want to do, or when I want to deceive my partner in an action and a little drop of lying substance corrodes the rim of my own heart, and I nonetheless do the evil, although 'not with the whole soul', what role has 'reason' played in the event? It was not at all, in fact, a thinking that took place there; it was only that gentle protest of the soul to which we so often are accustomed to pay no attention (Buber 1964, p 113).

Another quality concerns the unity of the self and is related to the normative aspect of the discussion concerning *Wesen*. Buber often cites the unity of the self as a never-ending but necessary task in order to become a more complete being. He goes so far as to claim that unity alone is real power and that only the unified person reigns (Buber 1962, p. 1047). This view is related to but not identical with the traditional conception of agency consisting in control over one's multiple volitional powers or the conception of freedom as mastery over one's actions. Freedom for Buber is not the freedom from interference but living in dialogue: "At the opposite pole of being compelled by destiny or nature or men does not stand being free of destiny or nature or men but to commune and to covenant with them. To do this, it is true that one must first have become independent; but this independence is a foot-bridge, not a dwelling-place" (Buber 1926, p. 108). Complete unity is impossible to achieve except in God (quoted in Reichert 1996, p. 37) but it can be strived after and will provide a better position from which to act morally.

Buber's assumption here is that unified actions can be brought about by unified agents, while agents with a more complex and conflicting "soul" are too restless to act in a way other than "patchwork" (Buber 1947, quoted in Reichert 1996, p. 176). The assumption that it is more promising for successful agency to be a unified agent is accompanied by the further conviction that it is possible to acquire such a unified self. Although people vary in their degree of self-unity – whether by nature or by grace – it is possible to work and live towards a more unified state of self:

The person with the multifaceted, complicated, conflicting soul is not abandoned: The innermost of this soul, God's power in its depths is capable of operating on it, to change it, to bind together the forces that are at feud with each other, to merge the elements that strive away from each other, to unify them. Such a unification must happen, before the person attempts an extraordinary piece of work" (Buber 1947, quoted in Reichert 1996, 176, translation B.S.).

Perfect unity is impossible to achieve for a human being but every piece of action carried out by the collected soul helps to unify and strengthen the self.⁴⁹

Buber's critical approach to psychological models that postulate one single or one leading drive is part of his pluralistic conception of the human being; "In opposition to these doctrines and methods, which impoverish the soul, we must continually point out that human inwardness [*Innerlichkeit*] is in origin a polyphony in which no voice can be 'reduced' to another, and in which the unity cannot be grasped analytically, but only heard in the present harmony" (Buber 1926, p. 102).

One further point should be emphasised. Although Buber relies heavily on the Chassidic culture for his references about the human being and the net of interconnections that contribute to her self, his anthropological philosophy builds on

⁴⁹ I have remarked before that Buber uses a number of concepts to denote the self, including the widely discredited 'soul'. His insistence should be remembered however, that even when he uses the word "soul", he means: "the whole human being, body and *Geist* together. The soul is not really unified, unless all bodily forces, all limbs of the body are" (Buber 1947, quoted in Reichert 1996, p. 177, translation B.S.).

more mainstream insights of the two centuries preceding his. Similar to the narcissistic insults that humanity received by the hands of Copernicus, Darwin and Freud, there are three philosophical insights behind which there is no possibility of a return. Buber postulates these to be the discovery of the historicity of being (G. Vico), the forms of intuition (I. Kant) and the sociologically and psychologically determined variety in the perception of individuals (modern analysis of society and the psyche; meaning first and foremost, we can presume, Marx and Kant) (quoted in Reichert 1996, p. 27). What these revolutions in the history of thought mean for Buber's conception of the nature of self and of morality is not explicitly stated. But it does suggest that the moral agent is not simply to be understood as a mythical creature who, when facing a situation, miraculously succeeds in answering its demands.

Buber considers two social factors, which have contributed to make the anthropological problem of the human more visible. He also argues that historical changes are also partly responsible for the new crisis that human beings find themselves in as a result of modernity. One reason why the philosophical problem of anthropology has become more vivid is the decrease in the importance of traditional and organic forms of social living like family, rural communities or work co-operatives; i.e:

communities which quantitatively must not be too big to allow the men who are connected by them to be brought together ever anew and set in a direct relation with one another, and which qualitatively are of such a nature that men are ever anew born into them or grow into them, who thus understand their membership not as a result of a free agreement with others but as their destiny and as a vital tradition" (Buber 1938, p. 1986).

Other forms of living and working together have established themselves; but such associations, unions or political parties, although coupled with passionate activity, cannot provide the human being with the sort of security that previous settings could

do (ibid..). Thus, for Buber it's the very freedom with which contemporary associations are sought and chosen that make such associations lack the element of a 'home', the crisis that Buber regards as characteristic for modernity as a "crisis of confidence" (Buber 1948, p. 233) which he considers to be directly linked with the lack of life in small, organic communities.

What leads the modern human being into a specifically modern crisis is the second factor that specifies the problem of philosophical anthropology. Buber calls this "peculiarity of the modern crisis man's lagging behind his works" (Buber 1938, p. 187). In three areas of modern development, Buber argues, in technology, in economy and in politics, modernity brought an increase in power and a connected increase in helplessness and loss of control (ibid.). Due to these two factors – the loss of organic relationships and being overtaken by human creation – the *Wesen* of individuals is in danger of missing the mark of what Buber considers to be necessary for the moral life, both in the sense of the good and the right life.

One aspect of Buber's recurring critique of the modern outlook on man concerns the analytic and reductive approach to self and consciousness. Buber's *Wesen* is characterised by an imperfect wholeness, unity and uniqueness (Buber 1954, p. 285). The view of the modern human being however, as is the postmodern being that succeeded him in a much more emphasised manner, is the product of an analytic, reductive and deductive approach. Part of Buber's critique seems to be directed at the vogue of psychoanalysis, which regards the bodily and psychic being of humans as a composite, which can be dissected and analysed. Buber opposes this by denying that the stream of consciousness itself can ever be grasped objectively (ibid.). The modern approach is reductive because it tends to reduce the complexity and diversity of the individual, which is "nourished by the microcosmic fullness of

the possible” (Buber 1954, p. 285, translation B.S.) to recurring schemata. What Buber criticises around the term “deductive” is less clear. It is related to the study of the development of humans, which Buber interprets as being performed around “genetic formula”, which disregards the “dynamic central individual-principle of this development” (Buber 1954, p. 285, translation B.S.).

Buber is critical of the role of reflection in the process of becoming a *Wesen* and does not accord it the place that it traditionally enjoys:

[The] emerging I becomes conscious of itself but without becoming an object for itself via reflection. [...] Naturally the presence of persons is necessary for a personal encounter to take place; but the degree of the development of I-consciousness or even the degree of its reflective elaboration is not an essential (*wesentlich*) moment. I see that Socrates reflects, I don’t see that Franciscus does; both their relations to this or that student are genuinely personal” (Buber 1963, 595, translation B.S.).

Part of Buber’s anthropological agenda, i.e., his study of the *Wesen* of the human being can be found in his discussion of the ‘principle’ (*Prinzip*) of being human (Buber 1950, p. 411). Like the duality in the attitudes that a human being can be in, this ”principle of being human” consists in a “double movement”. The primal distance (*Urdistanz*) – is the precondition of the other, the entering into an encounter (*In-Beziehungtreten*) (ibid..). Initially, Buber’s discussion of these two anthropological principles seems to run counter the double aspect of human beings, i.e., the existence of the I only as part of an I-It relation or I-You relation. Buber argues that the precondition of encountering an other is the latter’s existence independent from the subject’s perception and he argues moreover that this is only possible for human beings (Buber 1950, p. 412). The attitude of human beings to the world, postulates Buber, extends beyond what other animals experience as their environment and abstracts a unitary state of being (ibid..), which is the precondition for encounter.

The life of a human being runs in a discontinuity between the two modes of being in an I-You relation, which is characterised by authenticity (*Eigentlichkeit*) and the I-It relation, which is not (*Uneigentlichkeit*). This discontinuity may not be denied and cannot be abolished (Buber 1963, p. 637). Continuity in the realm of the I-You is impossible but by establishing a coherent set of possibly valid propositions, it might be possible to gain continuity in the realm of the I-It. However; “If one confines oneself to the “world”, one will suffer the inevitable loss of the authenticity (*Eigentlichkeit*) of being (*Dasein*) [...] that, within which and with which one has to live, is objectified in just that measure in which the transmutation is successful” (Buber 1963, p. 637, translation B.S.). Thus, an attempt at creating continuity in one’s life is bound to diminish the weight of the I-You realm. An alternative approach would be to accept the fact of discontinuity and allow the I-You realm more weight, even leadership of the I-You relation. Buber argues that the I-You can lead, although it does not provide security or guidelines but by continuing to exert its influence after the I-It relation has taken over again, “Also here, the discontinuity is not abolished ; we take it on us and we master it by the realised primacy of the dialogical” (Buber 1963, p. 638, translation B.S.)

For Buber character is an aspect of agency, which is developed with influence from the environment, which can be trained and which concerns her attitudes and actions. This concept should not be understood as identical with the more obscure *Wesen*. In an article, in which Buber discusses the “great character”, it can be interpreted as the ideal, to which education can aspire even though it is not possible to produce it as a natural product of education (Buber 1939a, p. 134). This idea, the great character

can be conceived neither as a system of maxims or as a system of habits. It is peculiar to him [the great character, B.S.] to act from the whole of his

substance. That is, it is peculiar to him to react in accordance with the uniqueness of every situation which challenges him as an active person. Of course there are all sorts of similarities in different situations; one can construct types of situations, one can always find to what section the particular situation belongs, and draw what is appropriate from the hoard of established maxims and habits, apply the appropriate maxim, bring into operation the appropriate habit. But what is untypical in the particular situation remains unnoticed and unanswered (Buber 1939a, 134).

The motifs that collect around the idea of the great character are thus the by now familiar ones that are to be found around the idea of the *Wesen*. The right response to the unique situation, argues Buber again, can only be produced by the individual who resists perceiving, judging and reacting to the given situation in a habitualised manner or according to a set of rules, which need only be applied to the actual moment. This does not turn Buber's great character into an existential figure who stands beyond norms however. As mentioned before, Buber does believe in moral norms that exist independently of humans and he believes in their prescriptive character, "But the command inherent in a genuine norm never becomes a maxim and the fulfilment of it never a habit" (Buber 1939a, p. 135).

Dialogue and Monologue

The figure of the agent then that Buber draws is characterised by being relational and the nature of the relation is characterised by being either in the realm of the I-It or the realm of the I-You. The ontological fact of being relational and the normative demand of responding to the address of the other establish the domain of dialogue and disparages that of monologue. These terms should not be understood in their narrowest sense however. Buber repeatedly remarks that this sort of dialogue or conversation should not be understood as exclusively vocal or written communication using signs and symbols although such is its proper domain (Buber

1930, . 144). Addressing and being addressed, responding and being responded to make up the domain of speech in its widest sense:

Human existence, even the most silent, is speech; and speech, whether intentionally or unintentionally, directly or indirectly, along with gaining ground and forcibly penetrating, along with sucking and tasting, along with advancing over untried ways, is always address. What addresses you, not in the said but in the saying, is the underivable person, the now living new creature (Buber 1964, p. 34).

Dialogue is not only valuable in itself but also as a means to an end. The end is related to Buber's anthropological views on the nature and 'purpose' of human beings. It is better rather than worse that a human being should develop in a way as to be as authentic as possible. The path that leads to authenticity does not consist of involvement with one's own self but with the world, encountering others in dialogue: "The series of meetings that a man has taken part in is more important for this personal existence than his total possession of impersonal scientific knowledge, no matter how highly this too is to be prized. It is the former that builds up the core of the person" (Buber 1964, p. 40). Conversely, a return to the objectification of the other leads to the dissolution of that core: "[...] the man who makes the other from a Thou into an It thereby destroys his own life at its core" (Buber 1964, p. 115).

Such overly lyrical expressions are typical of some of Buber's passages but the area of dialogue is nevertheless neither mystical nor loving, intimate or erotic, as some such descriptions might lead the reader to imagine. If communication does not take place in the medium of the spoken or written word, accessible to all, it loses its objectivity but it does not become "mystical" or finds its completion in such an event but rather "in one that is in the precise sense factual, thoroughly dovetailed into the common human world and the concrete time-sequence" (Buber 1930b, 5). Buber's postulation that it is an ontological fact that the I is relational leads to the dissolution of the I if the address of the other is ignored or degraded. But while drawing this

contrast between dialogue and monologue, he refrains from elevating moments of erotic love or similar, enthusiastic immersions in the other with what makes adds to what he calls the core of the person.

On the contrary, Buber emphasises that nowhere does the monological mingle with the dialogical as much as in the erotic (ibid..). He emphasises that the condition for the creation of a dialogue – or a “real conversation” – (Buber 1954, p. 283) is not, as might be assumed, agreement about the convictions of the parties or influencing the other with an aim to persuade her about a question at hand. The relation between the parties that might enable a dialogue can also be a fight, in which the personhood of the other is affirmed while her position is not (Buber 1954, p. 283). Friends, lovers or others can be as much de-personalised in a given moment of encounter as a stranger or enemy. It is not the title or the framework of the relationship that determines whether an encounter deserves the name of the dialogue but the openness and attentiveness of the participants towards each other: “Being, lived in dialogue, receives even in extreme dereliction a harsh and strengthening sense of reciprocity; being, lived in monologue, will not, even in the tenderest intimacy, grope out over the outlines of the self” (Buber 1930, p. 168 / Buber 2004, p. 24). The type of relation between two people, which it a dialogical one is the degree of *Umfassung* that constitutes it. A dialogical relation is thus one, which is characterised by “[...] the extension of one’s own concreteness, the fulfilment of the actual situation of life, the complete presence of the reality in which one participates” (Buber 1926, p. 116).

Buber further differentiates dialogical life from a dialogue and monological life from a monologue (Buber 1930, p. 168). Of three kinds of communication, only one is a genuine dialogue: “There is genuine dialogue – no matter whether spoken or

silent – where each of the participants really has in mind the other or others in their present and particular being and turns to them with the intention of establishing a living mutual relation between himself and them” (Buber 2004, 22). Buber calls such a “turning to” the other *Hinwendung* (literally “turn toward”) while the monological approach to the other is *Rückbiegung*⁵⁰ (literally “bending back”) as the monologist focuses on her own self in a reflexive bend. These types of basic action (*Grundbewegung*) could be understood as the units of a more general attitude, the *Wesenshaltung*. As so often, Buber’s meaning should be distinguished from an Aristotelian as well as a Kantian interpretation, since habitualisation as well as deliberation are not to be understood as part of the process (Buber 1930, p. 170/ Buber 2002, p. 25). The ordinary perception of the world is “an insignificant multiplicity of points to one of which we pay momentary attention” (Buber 1930, p. 170 / Buber 2002, p. 26). An act of *Hinwendung* concentrates the forces of the person, both physical and mental, on the other, an entity out in the world. An act of *Rückbiegung* on the other hand is not the act of the self averting itself but rather reflexion on the own I or a return to it. (Buber 1930, p. 170 / Buber 2002, p. 26). A return to the own I without averting from the other is Buber’s way of describing a common perception of the other, namely as “his own experience, only as a ‘part of myself’” (Buber 1930, p. 173 / Buber 2002, p. 27).

The ‘technical’ dialogue, which is merely a means to the end of exchanging information and the monologue that poses as a dialogue are examples of failed dialogue. The first of these failed dialogues, the technical one, is part of Buber’s critique of modernity. More interesting – and amusing – is Buber’s description of the

⁵⁰ In the English translation of Buber’s book *Zwiesprache*, this term is translated as “reflexion” (Buber 2002, p. 26). I prefer to use the original German term however, for although re-flexing is literally identical with as Rück-biegung, this old-fashioned spelling of the modern version ‘reflection’ is too heavily laden with other connotations.

“monologue disguised as a dialogue, in which two or more men, meeting in space, speak each with himself in strangely tortuous and circuitous ways” (Buber 1930, p. 166 / Buber 2004, p. 22). Examples of such monologues pretending to be dialogues are debates, which take the form of a competitive sport, conversations that are bent not on presenting or receiving information, building contact but are rather involved in strengthening and maintaining the self-conception of the participant(s), friendly conversations, in which the parties regard themselves as absolute and legitimate, the other as questionable and relative, lovers’ discourse, in which both the partner and the own “magnificent soul” is savoured (Buber 1930, p. 167).

Differences between the types of experiencing a given moment are further clarified when Buber differentiates among three ways of perceiving a human being in front of us. The observer (*Beobachter*) analyses as many of the singular parts of the physiognomy or movements; her aim is to draw as accurate a picture as possible. The viewer (*Betrachter*) allows his mind to drift and is only after an impression of what she watches. Buber believes that all great artists had this mode of perception (Buber 1930, p. 151). The third mode however does not regard the other as an object separate from oneself, which is to be studied and understood. Rather, the other “speaks to” the perceiver, not of herself but rather as a vehicle for the word and for being (Buber 1930, p. 153). Although living should be understood as being addressed by signs the tendency of the modern human is to remain in an armour, which keeps off signs (Buber 1930, p. 153).

The other whom I face in an encounter, in a real dialogue ought to be neither an object of my observation nor of my watching. Buber calls the third way of approaching the other *Innewerdung*, which enables such a genuine dialogue. *Innewerdung* is an old-fashioned word, which could be paraphrased as “suddenly

becoming aware (of someone or something)”. Buber’s discussion of this type of approach to the other is exemplary of the praise of unity or wholeness that he has a tendency to indulge in as well as his anthropocentrism. To become aware of a thing or *Wesen*⁵¹ is to experience it “as a whole but at the same time without reductive abstractions, in all concreteness” (Buber 1954, p. 284, translation B.S.). Humans are categorically different from all other things or *Wesens* because humans are endowed with *Geist*, which is one of their individuating and distinguishing characteristics (ibid.).

The description and prescription of dialogue that is in question in Buber’s writings should not be misunderstood as pertaining only to relationships between two or a few individuals. A dialogue is not only possible but also necessary between more collective and abstract entities such as peoples or nations. Accordingly, regarding the Israel-Palestine conflict, Buber spoke out for a common creation and joint venture of Jews and Arabs in one state (Buber 1939) and was a constant and vocal supporter of dialogue between the two peoples. However, since dialogue in the more narrow sense that Buber uses to depict encounters between two or at most a few beings in a particular situation is decidedly not the same sort of encounter and relationship that takes place between entities as abstract as nations or people, the usage of the word should not imply that the two senses of the word are identical. This might be one of the examples of carelessness with termini that make a correct interpretation of Buber difficult.

Buber’s plea for dialogue between two distinctly different entities reaches from marriage to the public sphere and involves a mundane, every-day otherness rather than a mystical, blurred concept that is sometimes attributed to him:

⁵¹ „*Wesen*“ in this context is nearer to its verbalised meaning of simply „being (a thing)“, rather than the totality of its attributes.

That the men with whom I am bound up in the body politic and with whom I have directly or indirectly to do, are essentially other than myself, that this one or that one does not have merely a different mind, or way of thinking or feeling, or a different conviction or attitude, but has also a different perception of the world, a different recognition and order of meaning, a different touch from the regions of existence, a different faith, a different soil: to affirm all this, to affirm it in the way of a creature, in the midst of the hard situations of conflict” (Buber 1930b, p. 72).

This political aspect of the value of dialogue is related to its function. Buber’s peculiar view on the world considers it as something that must be worked at to come to being. There is a potential of more or less authenticity in every creature and every encounter, so that dialogue is not simply an end in itself but it also serves to contribute to the most authentic state of being of, as Buber likes to call it, creation:

Affirmation could be misunderstood as something static. I encounter someone and affirm him, as he is now. But to affirm a person, as he is, is only the first step because affirmation doesn’t mean that I understand the appearance in this moment as the person whom I want to affirm. I must understand the other person in his dynamic existence, in his specific possibility (quoted in Reichert 1996, 72, translation B.S.).

Being in dialogue is therefore paradoxically a state, in which the focus on the other contributes to the creation of the self as a more authentic and unified entity. An authentic and stable self is produced not as a consciously and purposely pursued goal but as a side effect of practicing and perfecting the art of turning to the other, turning to the world. Buber’s championing a singular and particular path for each individual is thus ethical (Reichert 1996, p. 169) but it is ethical in a way that encompasses both the idea of the good life as aesthetic of existence as well as the choice of right action.

The Affirmation of the Other

Buber’s call for the attentiveness and open-mindedness that are needed to be able to respond to the demands of the situation are directly linked to his concept and

championing of the other. If one had to pick three terms to explain what Buber deems necessary for genuine dialogue, these would have to be attentiveness, open-mindedness and affirmation (*Anerkennung*). The affirmation (or recognition) in question here is an affirmation of the other. Not only of the other as another entity but specifically of the otherness of the other; of what makes it different, strange or even alien to the agent. The need for affirmation is peculiar to human beings; “the animal does not need to be affirmed, because it is, what it is, unquestioned” (Buber 1950, p. 423)

Some decades after Buber, the idea of recognition became wide-spread and widely discussed. I prefer to use the term “affirmation” instead of “recognition” in order to maintain the distinction between Buber’s connotation and the one it received in later years, especially in discussions around multiculturalism. The two areas are not unrelated. Thomas Hobbes had already remarked on the importance of “glory” as one of the three causes of trouble, elevating the need for recognition to the level of competition and diffidence as the main drives of human action and reasons of conflict (Hobbes 1996, p. 83). The thread continues with the familiar debate around the concepts of multiculturalism, communitarianism and recognition (Kymlicka 2002, Fraser & Honneth 2003, Honneth 1994). In this debate the communitarian side sets off to criticise the perceived atomism of individualistic social and moral theories and argues for the necessity of a more realistic picture of human beings, the being part of whom of a community is an integral part of their self. Recognition of the other therefore involves recognition of the principles and practices of the group that the other is a member of. Buber also takes it as a starting point that “The human person belongs, whether he wants to acknowledge it and take it seriously or not, to

the community in which he is born or which he has happened to get into” (Buber 1930b, p. 77).

Nevertheless, Buber positions himself against the idea that individual self-recognition or the recognition of the group one is a part of is essential in any way. He argues to the contrary, regarding these two types of recognition as illusory surrogates of that recognition which comes out of “direct and personal mutuality” (Buber 1952, p. 225). The illusion of self-confirmation is bound to end in a realisation of abandonment, while

confirmation through the collective [...] is pure fiction. It belongs to the nature of the collective, to be sure, that it accepts and employs each of its members as this particular individual, constituted and endowed in this particular way. But it cannot recognize anyone in his own being, and therefore independently of his usefulness for the collective (Buber 1952, p. 225).

Buber’s own position about the human being is much more particularistic than either the figure of the independent, atomistic individual or that of the individual as essentially part of a collective. It is this particular other in this particular situation to whom affirmation is due, not out of duty, reason, expediency or as a follower of liberal-tolerant political theory but because of Buber’s particularistic ontology and ethics, which regards the other as a constitutive element of both self and the situation.

Such an open approach to the other is far from being the common attitude, as Buber complains in the critique of his time. The demon that blocks the way of genuine dialogue to flourish among individuals and groups is mistrust (Buber 1952, p. 222). The question of trust and mistrust is directly related to the question of authenticity; it is the mark of what Buber considers to be “inherent in the human being” (Buber 1952, p. 223). What is being doubted here is the authenticity of the other; more specifically whether she means what she says and whether she would do

as she says (Buber 1952, p. 223). Buber argues that there has been a distinctly modern addition to this more common type of mistrust. This new aspect is the conviction that there is a discrepancy between word and deed or between word and meaning of the other party, not out of design but as what her *Wesen* dictates. Following the teachings of Marx and Freud,⁵² the modern man searches for hidden meanings and intentions behind the others' attitude and ideas are dismissed as ideology (ibid.).

Related to this diagnosis of his age, Buber postulates two main ways of influencing other people, their attitudes and behaviour (Buber 1954, p. 287); namely propaganda and education. The other for the propagandist is not fundamentally a person or an other but a carrier of the attitude or judgement that is to be increased and strengthened by propaganda. While individuality is important for a political party, not in itself but as a tool for the optimisation of functional specialisation, propaganda simply aims to increase of members and followers (Buber 1954, p. 288). He does not believe in the power of his stance to spread without help (Buber 1954, p. 289)

The educator on the other hand sees the other as a singular and unique set of potentialities, which are due to be actualised. While the propagandist aims to multiply his preferred modus of perception, evaluation and action in as many people as possible, the educator aims to support the individual in becoming what Buber calls the mission of being, which can only be realised by each particular individual (Buber 1954, p. 289). As so often in Buber, the picture he draws of the propagandist and the

⁵² Buber does not dismiss the teachings of Marx and Freud. What he objects to is rather the reduction of a great number of various factors that influence the human being to one or two newly discovered ones: "It is a matter of showing up a fundamental and enormously influential error of all the theories of seeing-through and unmasking. The gist of the error is this: when an element in the psychical and spiritual existence of man which formerly was not or was too little noticed is now uncovered or clarified, one identifies it with man's total structure instead of inserting it in this structure" (Buber 1952, p. 226).

educator are idealised caricatures, while real individuals combine these two principles in varying degrees (Buber 1954, p. 290); i.e., their approach to the other includes elements of regarding them as carriers of right action or conversely as actualisers of their own, individual mode of being.

What this diagnosis implies is more than a plea for more authenticity. It also involves a strong demand towards truthful expression. In the style of relationships, which Buber criticises, “[o]ne is still inclined to spare the other in order that one may oneself be spared” (Buber 1952, p. 224). We might call the position criticised here a comfortable quasi-liberalism, which bases its plea for laissez-faire not on a philosophical principle but on the expediency and the advantages of mutual non-intervention. Buber’s call however is radical in the dictum that one is obliged to disclose herself. He quotes a Zen-master whom he reports to have said “Speech is disdain but silence is deceit. Beyond speech and silence lies a steep path” (quoted in Reichert 1996, p. 271).

The centrality of the other is linked to Buber’s ontology of the I as a relational entity that can only exist as part of a pair I-It or I-You. The centrality of this claim for Buber is explicit in the following quotation:

Were I to tell a questioner, what should be understood the main, verbally conveyable result of my experiences and reflection, then there is no other possible response open to me than to avow myself to the questioner and the knowledge that enfolds me: To be a human being means, to be the *Wesen* that exists facing one⁵³ (Buber 1986, p. 83, translation B.S.).

⁵³ The original German text is: „Soll ich einem Fragenden Auskunft geben, welches denn das in gedanklicher Sprache aussagbare Hauptergebnis meiner Erfahrungen und Betrachtungen sei, dann ist mir keine andere Erwiderung gegeben, als mich zu dem den Fragenden und mich umfassenden Wissen zu bekennen: Mensch sein heißt, das gegenüber seiende *Wesen* sein“ (Buber 1986, p. 83). The last sentence, which is arguably the central sentence of Buber’s thought is difficult to translate while keeping the connotations involved: “Gegenüber” can be translated as “opposite one” as well as “facing one” and it includes the sense of remaining alien to what one faces.

Authenticity

Authenticity is arguably the culmination of Buber's line of thought, which we have been following so far. His ontological assumption that the self is relational brings the other into the forefront of ethical interest. Depending on the nature of the relation between the I and the other, the individual can develop into what Buber calls a person, who is capable of relating to the other in a dialogue and thereby contributes to the coming to being of an authentic I-You, in which she participates with all her *Wesen*. Many factors may contribute to a failure in this process and among these are the lifestyles, which are peculiar to modernity. A failure involves functionalising and objectifying the other, which is not only a morally blameworthy approach to him/her/it but also leads to a lack in self-realisation. Authenticity thus appears as the last of the central terms in Buber's thoughts and denotes a goal in Buber's teleological conception of the coming-to-being of the world and additionally a means to the end of achieving that goal.

In his discussion of authenticity Buber makes a distinction between being (*Sein*) and appearance or seeming to be (*Scheinen*), which he diagnoses as the real problem between individuals (Buber 1954, p. 277). Two poles of human existence correspond to this distinction: Life as lived from the *Wesen*, as one is and life lived from appearance, as one wants to appear (ibid.). Buber therefore understands "Truth" as being realised in an encounter, in which the parties show themselves as what they are (Buber 1954, p. 279). An encounter between two so-called *Bildmenschen* on the other hand, i.e., between two people who focus on the picture they draw for the rest of the world to see, turns out to be the meeting of six ghost-like figures rather than two persons. Buber asks us to imagine two such *Bildmenschen* in

a conversation – Peter and Paul – and counts the figures involved. First, there is Peter, as he wants to appear to Paul, then Paul, as he wants to appear to Peter; then Peter, as he really does appear to Paul, i.e., the picture that Paul has of Peter and conversely, the picture that Peter has of Paul. Then there is Peter as he appears to himself and Paul as he appears to himself. Finally there is Peter in flesh and blood and Paul in flesh and blood: “Two living beings and six ghostly appearances, which mingle in the conversation of the two in various guises” (Buber 1954, p. 279. Translation B.S.). Such a constellation, Buber adds, does not allow there to be a real encounter between two human beings.

Authenticity also demands that one express himself to the other by bringing his whole *Wesen* into the encounter: “And this means that he must be willing to say respectively what he has in mind about the subject talked about” (Buber 1954, p. 294, translation B.S.) Buber sees the need for such a self-declaration in his understanding of the In-between as an ontological sphere that can only come to being if the participants allow it. Accordingly, wrongdoing is to keep back; both from the situation and from the others by keeping back what one has to say. The appearance that such a stance brings into being keeps it from being authentic.

Philosophical Anthropology as Anthropocentrism

As mentioned before, Buber considers the area of much of his work to be philosophical anthropology. He reminds us that Kant’s famous questions considering what one can know, what one ought to do and what one may hope for, all refer to the fourth question concerning the nature of man. Kant thereby poses the main question of philosophy as anthropology but does not, according to Buber, live up to what he

promises. What lacks in Kant's discussion of the nature of human beings are the questions that are central for Buber:

[...] the question, what man is, is simply not raised, and not one of the problems which are implicitly set us at the same time by this question – such as man's special place in the cosmos, his connexion with destiny, his relation to the world of things, his understanding of his fellowmen, his existence as a being that knows it must die, his attitude in all the ordinary and extraordinary encounters with the mystery with which his life is shot through, and so on – not one of these problems is seriously touched upon. The wholeness of man does not enter into this anthropology (Buber 1938, 142).

A common strategy to approach the question of philosophical anthropology is the attempt to delineate human beings from other creatures. Searching for and postulating characteristics, which are distinctly human and which distinguish every single specimen of human beings from every other specimen of other animals are regularly attempted. Buber, although he is not human-centred when it comes to the question with which beings one can have an I-You relation, does like to claim now and then that there is something special about humans, not as a set of qualities that distinguish them from other animals but rather as a special way of being. This way of being is based on the ability of the human being to distance herself from the world of which she is part of in such a way as to separate distinctly subject from object.

This ability, paradoxically, makes the human being capable of being in a relation, while the animal remains ignorant of the state of being in a relation. Buber sees the reason for this to be in the impossibility of being in a relation with something, which is not regarded as separate and existing independently; separation between the two parties is a pre-condition to be in a relation (Buber 1950, p. 414). As remarked before that Buber does not understand this ability of distancing oneself from the other (*Urdistanzierung*) as a reflective one; it is not the higher cognitive

skills that turn humans to “the *Wesen*, through whose being the existing thing is set apart from it” (Buber 1963, p. 594).

“The gorilla too is an individual, a termitary too, is a collective, but I and Thou exist only in our world, because man exists, and the I, moreover, exists only through the relation to the Thou” (Buber 2002, p. 243).⁵⁴ Not only is the I-You relation proper to humans, it is proper to relations between humans;

The fundamental fact of human existence is the human being with the human being. What distinguishes the human world in its singularity is above all that something takes place here between *Wesen* and *Wesen*, nothing similar to which can be found anywhere else in nature (Buber 1948, p. 404).

What is confusing about such passages is not only the anthropocentrism behind it, which seems to deny I-You relations between humans and other animals or among non-human animals. What is more disconcerting is that the sentiments displayed here contradict Buber’s assurances of I-You relations being possible between humans and art, animals or nature. There is another interpretation, which might save Buber from the offence of inconsistency, if he regards the peculiar I-You relations among humans as one type and as the peculiarly human type of encounter. Understood as such, the inconsistency vanishes but leaves Buber with the task to explain what it is about the human-human encounter that elevates it above all other encounters in the world.

Another instance of Buber’s anthropocentrism is displayed in his discussion of the importance of affirmation. The reason why affirmation is crucial for humans is that it is the tool of turning potentiality into actuality, which only human beings possess:

Because man is the sole creature known to us in whom the category of possibility is so to speak embodied, and whose reality is incessantly enveloped

⁵⁴ “Auch der Gorilla ist ein Individuum, auch der Termitenstaat ist ein Kollektiv, aber Ich und Du gibt es in unserer Welt nur, weil es den Menschen gibt, und zwar das Ich erst vom Verhältnis zum Du aus.” (Buber 1948, p. 407)

by possibilities, he alone amongst them needs confirmation. Every animal is fixed in its this-being, its modifications are preordained, and when it changes into a caterpillar and into a chrysalis its very metamorphosis is a boundary; in everything together it remains exactly what it is, therefore it can need no confirmation; it would indeed be an absurdity for someone to say to it, or for it to say to itself: You may be what you are. Man as man is an audacity of life, undetermined and unfixed; he therefore requires confirmation, and he can naturally only receive this as individual man, in that others and he himself confirm him in his being-this-man (Buber 1953, p. 136)

The Place of Divinity in Buber's Thought

Martin Buber's unusual attitude towards religion,⁵⁵ divinity and Judaism might partly have its roots in the diversity and variety of influences in his early childhood and youth. His parents having separated when he was very young, Buber lived with his grandparents, both of whom were intellectually and religiously stimulating (Buber 1986). In his early youth Buber distanced himself from the Jewish tradition and turned to philosophy, focusing and being especially influenced by Nietzsche and Kant. The revival of his interest in Judaism in general and the Chassidic culture in particular in his later life should not be understood as a return to the belief system of his childhood but rather as the creation of a singular and uncommon understanding of God, creation and their place in human life. This is one reason why it might not be possible to dismiss Buber's thought as religious in general or Jewish in particular. The question remains however, how fundamental divinity and God is for Buber's thought and whether it is possible to interpret it in a way as to make it accessible to an atheistic or agnostic readership.

⁵⁵ An example for Buber's wariness towards religion can be found in the following quotation: "Whenever there was once again religion in history, there was also a power in it, which – not in a contestable way like profane powers, but with highly legitimate appearance – distracted from God. The reason why it was very successful with this, is mostly the fact that it is much more leisurely to deal with religion than with God, who sends one out into restless wanderings from home and paternal house" (Buber 1963, p. 636, translation B.S.)

The main question concerning God and the divine in Buber's thinking is therefore whether they are dispensable or not. Thomas Reichert (1996) supports the view that Buber is also an author for non-religious readers, since the *conditio humana* that underlies Buber's philosophy is universal (p. 9). Maurice Friedman distinguishes between God as the source of prescriptions in Buber's philosophy of religion and the concept of genuine/authentic existence as the source of prescriptions in his philosophical anthropology (Friedman 1963, p. 153). As the 'In-between' between man and man is the focal point of Buber's investigation, Friedman argues that Buber's ethics, which is developed from his philosophical anthropology can be understood secularly, while his ethics, which is developed from his religious philosophy is distinctly not capable of functioning without the idea of divine revelation (Friedman 1963, p. 154). Distinguishing in this way between moral aspects in Buber's thoughts that have their foundation in divinity and such aspects that are founded in the space of 'In-between' of also non-religious encounters, can enable secularists to benefit from his moral philosophy without having to throw out the baby with the bath-water.

Friedman's claim that the sphere of morality and religion in Buber's thought can be neatly separated is open to doubt. Friedman argues that the real distinguishing border for Buber is not between religion and morals "but between religion and morals that remain attached to the general and religion and morals that remain attached to the concrete" (Friedman 1963, p. 161). Buber himself comments on the relationship between religion and morality in a characteristically contradictory manner. On the one hand, contrary to Friedman's claim he asserts that the sphere of religion and that of morality are not separate: "What concerns me fundamentally is that our relation to our fellow man and our relation to God belong together, that their

basic character, that of a reciprocal I –Thou relation, joins them to each other; practically speaking, that in reality there does not exist a special sphere of ‘religion’ and a special sphere of ‘ethics’” (Buber 1964, p. 44). On the other hand he postulates that there is a distinguishing element between the two, which gives religion more concreteness as well as more dignity than ethics:

Religion, certainly, has this advantage over morality, that it is a phenomenon and not a postulate and further that it is able to include composure as well as determination. The reality of morality, the demand of the demander, has a place in religion, but the reality of religion, the unconditioned being of the demander, has no place in morality. Nevertheless, when religion does itself justice and asserts itself, it is much more dubious than morality, just because it is more actual and inclusive (Buber 1930b, p. 21).

Buber description of his work as philosophical anthropology and his rejection of the label of theology seems to support Reichert’s view that he can be interpreted secularly:

Certainly, when I try to explain the fact of human being, I can never disregard that he, the human being, lives facing God; but I cannot involve God himself in my explanation in any point, just as I cannot abstract the for me indubitable operation of God in history from it and make it to the object of my examination. I do not recognise any theological anthropology in this sense, as I do not recognise any theological world history – I only recognise a philosophical one. [...] I am bound not only to the philosophical language but also to the philosophical method” (Buber 1963, p. 590, translation B.S.).

In this passage Buber is very clear both about the fact that he himself believes in the existence in God and moreover believes that an anthropological study of human beings must involve this aspect. He is just as clear about the fact however, that this study can be or even must be done by remaining agnostic about God. Passages like this imply that Buber for an atheist or agnostic interpretation of his work is possible.

Not all of Buber’s remarks support this independence of God however. One question about the place of God or the divine in Buber’s thought is especially important for its proper evaluation. If one dispenses with his position, shared with the

Chassidic conception of man and his place in the universe, that both are creations and that every individual has a special part to play in the manifestation of God's creation, how maintainable is Buber's 'system'? If it is true that "You need God in order to be, and God needs you – for that which is the meaning of your life" (Buber 1970, p. 130), how is right action – in accordance with what the "meaning of your life" is – possible without such a theoretical reliance on God?

The aspect of the place of God or of the divine in Buber's philosophy in general and moral philosophy in particular concerns the age-old question, whether a person can be a moral agent if she does not believe in God. This question is not identical to the meta-ethical question if there can be moral values unless they are formulated by God but can be interpreted as related to it. Buber's answer is unequivocal while answering to a critic who understands him to be teaching that no action can be right unless it originates in a bond with God: "If I taught that I would have to believe that a person who does not believe in God (or thinks that he does not believe in him) could not act morally. I am by no means of that opinion" (Buber 1963, p. 596). Once again it is not God but the absolute character of morality that Buber regards as essential and a moral action is "accessible also to the autonomy that understands itself as godless" (Buber 1963, p. 599) – "we talk, not of God but of the encounter" (Buber 1963, p. 600).

The necessary link between morality and God or the absolute is Buber's belief that moral values exist independently of human beings and that they are binding. It is possible to argue however, that Buber's understanding of God as the absolute provides a way of bypassing a theist fundament for Buberian ethics. Interpreted in this way, God in Buber is nothing but a reformulation and a short-cut to moral realism and absolutism. Even if Buber's intentions would be undermined by

this interpretation, it would not lessen the value of his moral phenomenology. However, all in all I believe that such an interpretation is not as far from Buber as this would imply. Buber was very wary about religion and the image of God painted by them; he seemed to think that the commandments and prescriptions in holy books were human in origin, not divine (Fox 1963, p. 141).

Buber's personal attitude towards God exemplifies this point. When attempting to point to the similarities between G.E.Moore's account of friendship and Buber's I-You, Putnam writes that while the atheist Moore's ideal friendship is "simply a good thing [...] for Buber the 'I-You' relation to the friend points beyond the friendship, points to and ideally leads to the relation to the divine, to the ultimate You" (Putnam 2008, p. 65). Nevertheless, Buber's personal belief in God is atypical and not to be taken as the literal belief that something like God exists. He writes that he does not believe in God, if believing in God means to be able to speak of him in the 3rd person. His belief in God is as someone or something *to* whom he can speak, not *of* whom (Buber 1986, 5p. 6). Analogous to his position towards morality, he rejects a notion of religion, which stands over and above the every-day life. His attitude towards religion and morality both reject the idea of special spheres and instead recognises nothing but the "demand of the hour and responsibility" (Buber 1986, p. 60, translation B.S.)

Buber's Moral Philosophy

In the preceding pages I have drawn a brief framework of Buber's understanding of the relational nature of human beings, the duality of their being and the qualities of being in an I-It realm as opposed to an I-You realm. I now turn to the question of

what implications his views of the nature of the self have for morality. What, in other words, does Buber's moral theory involve and how is it related to his anthropological premises? How can ethics incorporate Buberian ideas of morality that defy traditional conceptualisations and have been termed "Ethics without Norms" or "Lifeorientation without direction" (Reichert 1996, p. 9)?

If we understand the purpose of moral philosophy to find answers to the question of what should be done, then Buber does not only provide no answers himself but he argues that there are no answers to give, no universal system of ethics (Buber 1963, p. 616). The question, what one should do, is an empty question, "One is not to do anything. One cannot help himself, with one there is nothing to begin, with one it is all over. He who contents himself with explaining or asking or discussing what one is to do talks and lives in a vacuum" (Buber 1919, p. 109) The focus of critique here is on "one", on the notion that there could be universal declarations about what had better be done in particular situations. It should not be misunderstood as meaning that there are no absolute values.

Buber is both a realist and an absolutist when it comes to moral values. He believes in an intrinsic worth of moral values:

As the ethical in this strict sense we understand the Yes and No of the human being to the attitudes and actions possible for him, the radical distinction among them, which affirms or rejects these in accordance with this radicality not according to the benefit or harm for individuals and society, but according to the value or un-value that is intrinsic to them⁵⁶ (Buber 1953a, p. 575).

Buber thus rejects both relativism, the idea that moral values exist by the creation of particular groups and subjectivism, the idea that the individual is the originator of

⁵⁶ The original of this passage is: „Unter dem Ethischen in diesem strengen Sinn verstehen wir das Ja und Nein des Menschen zu den ihm möglichen Haltungen und Handlungen, die radikale Unterscheidung zwischen ihnen, die sie dieser Radikalität gemäß nicht nach ihrem Nutzen und Schaden für Individuen und Gesellschaften, sondern nach dem ihnen selber innewohnenden Wert und Unwert bejaht und verneint.“

moral values (Buber 1939a, p. 128). He believes that this sort of scepticism is characteristic of modernity and that it is “due to the disposition of a dominant human type of our age” (Buber 1939a, p. 130). As mentioned before, Buber sees the decline in traditional forms of living together and their replacement by collectivities, which provide a sense of identity and provide the norms that the agent regards as the highest instance as one main reason why this sort of scepticism is characteristic of modernity (Buber 1939a, p. 131).

Buber is convinced that in a given situation there is an absolutely correct response. Additionally, he takes it to be open to every individual to believe prescriptions to be absolute and use them as guidelines (Buber 1963, p. 616). What he doubts is not the absolute truth of a prescription, for example of the biblical dictum “honour your father and your mother” but rather what this fairly abstract piece of prescription is supposed to mean in the actuality of a given situation (*ibid.*), with particular children and particular parents. The task of the moral agent is difficult; she must not only be critical towards prescriptions with a claim to absoluteness – whether based on God or moral truths – but she must additionally be able to withstand the temptation of literal or traditional interpretations and be ready to interpret anew every possible principle in every single case.

Sorabji reports of a similar stance that Stoics and Aristotles shared with regard to rules and their role in moral decision making. He interprets both as having dismissed the idea that “ethics could be covered by a system of rules” and instead espoused the view that “perceptivity is needed that can bend rules, and that if one lacks that perceptivity oneself, one must follow someone who has it. As with Aristotle, the wise person is the only interpreter of and standard of what is lawful” (Sorabji 2006, p. 164). Although this position is similar to that of Buber, there is a

discrepancy in the idea of a mentor, the virtuous person whom one can turn to and emulate. It is the problem of deciding who is wise. A person who feels unwell might choose to go to a doctor but the choice of which type of doctor to go already contains decisions made on the subject. Similarly, the choice to decide which allegedly virtuous person – or more common today, which opinion leader – one should turn to already contains some idea of what one is likely to hear, which in turn arguably corresponds to what one wants to hear.

As demonstrated in the exposition of Buber's brand of situationism, a central concept of Buber's moral theory is responsibility. The concept is directly related to that of responding; or as the German word of *Verantwortung* implies to be in a state of answering a demand, a duty or a call. Due to the centrality of responsibility – the ability and the need to respond – Buber's moral theory cannot be positioned using the conventional taxonomy of moral theories. The criterion of morality – or in Buber's words, of the capability of responding – is attentiveness. Nothing more than attentiveness is needed, according to Buber, to begin interpreting the sign that the situation provides (Buber 1930, p. 161). None of the more traditional contestants of characteristics, which are needed to make one moral is evoked here. Neither rationality, nor experience, consistency or practical wisdom but solely attentiveness is needed to make the human being capable of responding to a given situation.

The trust that Buber thus displays with the human being is reminiscent of Socrates' position about success and failure in morality. It implies that given a genuine attention to the characteristics of the situation, a human being will be capable of generating a more or less appropriate response. As Buber insists the response will never be the ideal response but the reason for this is the high standard he sets for what an ideal response would be. No human being can respond ideally to

a situation because every human being is fallible and incomplete. But every human being, as long as she is attentive, is capable of generating a morally acceptable attitude towards a situation. Similar to Socrates' candidate for moral acceptability then, ignorance – in this case literally ignoring the situation – is the reason for moral failure.

Aristotle is a prominent critic of this implicit trust in the human capacity of arriving at moral truths independently of age, experience or wisdom. His concept of *phronesis* (Aristotle 1996) aims to highlight the importance of having both the necessary experience to judge a situation accurately and of having a good idea about the causes or possible consequences of a situation. Such skills are acquired gradually and by experience.⁵⁷ No young person – the famous “clever adolescent” of contemporary virtue theory – is capable of reacting appropriately to a situation in the same way as a person who has acquired *phronesis* over the years. This Aristotelian criticism of the Socratic confidence in knowledge or its adaptation to Buber would reject the idea that attentiveness to a situation is all that is needed to correctly perceive and interpret the signs.

Such an adaptation to Buber would miss the mark however. Although elements of dialogue such as attentiveness, open-mindedness and responsibility are the main factors that enable a person to act as a moral agent, an aspect comparable to

⁵⁷ An older, more discerning Keynes provides an example. With hindsight, he writes of their youthful ignoring of Moore's chapter „Ethics in Relation to Conduct“ concerned, not with goodness as a quality of states of mind but with rightness as a quality of action in a time when the Bloomsbury Group enthusiastically espoused Moore's trinity of beauty, friendship and truth: “We entirely repudiated a personal liability on us to obey general rules. We claimed the right to judge every individual case on its merits and the wisdom, experience and self-control to do so successfully” (Keynes 1972, p. 446). One problem with this approach is the fallibility of humans in general and the challenges posed to situationism in particular. Keynes: “What matters a great deal more is the fact that it [the aspect of our code] was flimsily based, as I now think, on an a priori view of what human nature is like, both other people's and our own, which was disastrously mistaken” (Keynes 1972, p. 447).

phronesis is also present in Buber's moral philosophy. This is grounded in the fact that the person who goes out of a dialogue or an I-You moment is not the same person as before; they have traversed ground in becoming more authentic, more unified, more in unison with their *Wesen*. Time and experience add to the 'quality' of moral responses, in Buber as well as in Aristotle. The difference is that while in Aristoteles habitualisation is an important aspect of becoming and doing good, in Buber habitualisation, i.e., acting in patterns according to abstract ideas of good, is the very death of real moral agency. The moral response in a situation must be authentic and spontaneous, but the spontaneity is the product of responding with the whole *Wesen*, and as such it "incorporates everything that one once was, including the own past I-You relationships with this person and with others" (Friedman 1963, p. 171).

Buber considers a question concerning his prescriptions⁵⁸ towards self-reflection or contemplation (*Besinnung*), choosing one's own, particular path and bringing the self, the *Wesen* to as complete a unity as possible. His answer demonstrates once again that his stance is fundamentally ethical: "Not for my own sake. [...] To begin with oneself, but not to end with oneself; to go out from oneself, but not to aim at oneself; to grasp oneself, but not to be intent on oneself" (Buber 1947, quoted in Reichert 1996, p. 185). This dictum is partly to be explained with Buber's espousal of the Chassidic notion that the particular path, which each individual ought to choose is part of the creation of the world. The ultimate goal is not the own good of the individual but serving this 'master-plan' of creation. Reichert, who sets out to demonstrate that Buber can be interpreted secularly, "for

⁵⁸ Without doubt, Buber would flinch at the term 'prescription'. Nevertheless, I think this is the proper term to use, since Buber does not always seem to follow his own dictums.

atheists” as well,⁵⁹ argues that realistically, every action and omission affects being and is not simply a non-binding event in the self (Reichert 1996, p. 186). And indeed, another passage by Buber could be interpreted as reformulating the image of creation as “human order of being” (*menschliche Seinsordnung*), in which every person is in objective relations to others, the totality of which constitutes her contribution to the state of being of the world and towards which she is responsible (Buber 1957, p. 486).

On Good and Evil

Buber’s approach on good and evil mirrors his approach to moral philosophy in general: “I was concerned above all to show that in their anthropological reality, that is, in the factual context of the life of the human person, good and evil are not, as they are usually thought to be, two structurally similar qualities situated at opposite poles, but two qualities of totally different structure” (Buber 1953, p. 64). This passage is characteristic because it demonstrates two of Buber’s main moral concerns: It asks for the problem of moral value to be bound with “anthropological reality” and it differentiates the two moral aspects under scrutiny as different in kind rather than in degree.

In his discussion of good and evil, Buber follows the Judaic tradition, which he interprets as regarding the “evil ‘urge’ as passion, that is, the power peculiar to man, without which he can neither beget nor bring forth, but which, left to itself, remains without direction and leads astray, and the ‘good urge’ as pure direction, in other words, as an unconditional direction, that towards God” (Buber 1953, p. 97).

⁵⁹ Thomas Reichert’s edition of excerpts by Martin Buber is correspondingly titled “Buber for Atheists” (Reichert 1996).

Both drives are necessary for agency in general and moral agency in particular. Similar to the evaluation of the realm of the I-It, the realm of evil is not to be dismissed or battled against. What is necessary for successful (moral) agency is to use the drive, which Buber interprets evil to be with the direction that good may provide: “To unite the two urges implies: to equip the absolute potency of passion with the one direction that renders it capable of great love and of great service. Thus and not otherwise can man become whole” (Buber 1953, p. 97).

In his discussion of good and evil Buber considers a variety of conceptions in the history of thought. He demonstrates that the idea of evil as a negative concept, i.e., as one, which is determined by the lack of direction rather than positively as a particular direction is not peculiar to Judaism. In the Persian mythology that Buber discusses, the God Zurvan “[...] does not choose, he doubts. Doubt is unchoice, indecision. Out of it arises evil” (Buber 1953, p. 104). Other motifs of the Iranian scriptures also seem to echo Buber’s own position. He reports of elements of authenticity, which are related to knowledge; “[...] doubt of being is the evil, the good is ‘knowledge’, belief in being [...] Here it is ultimately a question of fidelity and infidelity to being” (Buber 1953, p. 104) while another quotation from the Avesta concerns the question of awareness: “‘All good thoughts, all good words, all good deeds, I do consciously. All evil thoughts, all evil words, all evil deeds, I do unconsciously’” (quoted in Buber 1953, p. 106).

This dichotomy of direction and lack of direction is mirrored in the development of the person. Similar to Erik Erikson, who considered developing an identity as the developmental challenge proper to adolescence, in which the individual is faced with a multitude of possibilities, Buber argues that “the evolving human person [...] is bowled over by possibility as an infinitude” (Buber 1953, p. 125).

Such periods, generally to be found during puberty are periods of chaos, of the biblical *tohuwabohu* (ibid.). Referring to Piaget's model of cognitive development, Baird also describes the first phase of adolescence as one in which "[...] the increased ability to think hypothetically produces unconstrained thoughts with unlimited possibilities" (Baird 2008, p. 326). The danger of being seduced by such a vista of possibilities can also keep erotic encounters as monologues: "Many celebrated ecstasies of love are nothing but the lover's delight in the possibilities of his own person which are actualized in unexpected fullness" (Buber 1930b, p. 5)

In short, there are two types of state, in which the agent can find herself: One is characterised by a lack of direction and of decision; the *Wesen* is not collected, rather different motivational centres of the self strive towards different aims or are altogether aimless. This state, Buber associates with "evil" – evil is not the product of a decision to perform an action, which is considered to be the opposite of good but rather the lack of action as a result of indecision. Agency is hindered by indecisiveness and it is this type of non-action that is 'evil' in the terminology that Buber uses. Conversely, when one can speak of a genuine decision, i.e., a decision that is not under the authorship of a part but rather of the "whole of the soul" (Buber 1953, p. 130), we can also speak of 'good'. Buber's much quoted slogan that summarises this unusual approach to good and evil is: "Evil cannot be done with the whole soul; good can only be done with the whole soul" (Buber 1953, p. 130). Although in many places Buber discusses evil as the lack of direction and patchwork action without involving the whole *Wesen*, there is one positive characteristic of evil. It is the lie as the specifically evil, which has been brought to nature by human beings (Fox 1963, p. 147).

Before turning to an evaluation of the two aspects of Buber's thought that are relevant for my purpose, i.e., to his understanding of personhood and his moral theory, it is worth quoting a longer passage in full since it summarises the main aspects of the position of the moral agent in a given moral situation succinctly: "The words of our response are spoken in the speech, untranslatable like the address, of doing and letting – whereby the doing may behave like a letting and the letting like a doing. What we say in this way with the being is our entering upon the situation, into the situation, which has at this moment stepped up to us, whose appearance we did not and could not know, for its like has not yet been. Nor are we now finished with it, we have to give up that expectation: a situation of which we have become aware is never finished with, but we subdue it into the substance of lived life. Only then, true to the moment, do we experience a life that is something other than a sum of moments. We respond to the moment, but at the same time we respond on its behalf, we answer for it. A newly-created concrete reality has been laid in our arms; we answer for it. A dog has looked at you, you answer for its glance, a child has clutched your hand, you answer for its touch, a host of men moves about you, you answer for their need" (Buber 1936, p. 163).

Demands and Perfectionism

A common point of criticism of Buber's thought is its "impossibly demanding" nature⁶⁰ (Putnam 2008, p. 59). Putnam argues that "The famous 'I-Thou' in Buber is a relation that Buber believes is demanded of us, and without which no system of

⁶⁰ Putnam describes Buber as a moral perfectionist, using a term that he borrows from Stanley Cavell; "Such a philosopher is a 'perfectionist' because she or he always describes the commitment we ought to have in ways that seem impossibly demanding" (Putnam 2008, 59).

moral rules and no institution can have any real value” (Putnam 2008, p. 60). It is true that Buber’s ethics is demanding and it is true that a great deal is lost if an agent is not capable of being in an ‘I-You’ relationship. However I argue that Putnam’s criticism, representative of a common interpretation of Buber, mistakes Buber’s position. The I-You relation is not something that has to be acquired by arduous moral (self-)education and exercising cognitive and emotional faculties, by acquiring experience and wisdom. This becomes obvious if we bring to mind some of the examples of the areas in which an agent is in an I-Thou relation and what sort of agents are capable of it. Typical examples of an I-Thou relation are children engrossed in play or – as Buber’s own biographical story goes⁶¹ – in contact with an animal. Although the transition into an I-Thou relation can happen with an “illumination” or “awakening”, such examples are more the exception than the rule; “I already find this relation – as I have maintained from the beginning – in the life of the small child, as in that of the so-called primitive man, in a directly natural form; and I also understand the meaning of most spiritual forms in their connection with the natural” (Buber 1964, p. 31).

⁶¹ Buber’s own story of an encounter with a horse and his eventual relapse into what he calls *Rückbiegung* deserves to be quoted in full: “When I was eleven years of age, spending the summer on my grandparents’ estate, I used, as often as I could do it unobserved, to steal into the stable and gently stroke the neck of my darling, a broad dapple-grey horse. It was not a casual delight but a great, certainly friendly, but also deeply stirring happening. If I am to explain it now, beginning from the still very fresh memory of my hand, I must say that what I experienced in touch with the animal was the Other, the immense otherness of the Other, which, however, did not remain strange like the otherness of the ox and the ram, but rather let me draw near and touch it. When I stroked the mighty mane, sometimes marvellously smoothcombed, at other times just as astonishingly wild, and felt the life beneath my hand, it was as though the element of vitality itself bordered on my skin, something that was not I, was certainly not akin to me, palpably the other, not just another, really the Other itself; and yet it let me approach, confided itself to me, placed itself elementally in the relation of *Thou* and *Thou* with me. The horse, even when I had not begun by pouring oats for him into the manger, very gently raised his massive head, ears flicking, then snorted quietly, as a conspirator gives a signal meant to be recognizable only by his fellow-conspirator; and I was approved. But once—I do not know what came over the child, at any rate it was childlike enough—it struck me about the stroking, what fun it gave me, and suddenly I became conscious of my hand. The game went on as before, but something had changed, it was no longer the same thing. And the next day, after giving him a rich feed, when I stroked my friend’s head he did not raise his head. A few years later, when I thought back to the incident, I no longer supposed that the animal had noticed my defection. But at the time I considered myself judged.” (Buber 1930, p. 171 / Buber 2002, p. 26).

Children and animals are also important partners in encounters with adults and are capable of challenging, changing and “educating” them (Buber 1995, p. 20 / Buber 2004, p. 20). Art and nature are other areas, in which an agent finds herself in a dialogue with the other. A typical example of an I-You relation is decidedly not an elderly professor in contemplation of a captivating problem. It is also not the monk lost in meditation; “All the prescriptions that have been excogitated and reinvented in the ages of the human spirit, all the preparations, exercises and meditations that have been suggested here have nothing to do with the primally simple fact of encounter” (Buber 1970, p. 125). The reason why the monk immersed in meditation is not a good example for the elements of the encounter in Buber’s sense is that for him, the encounter is not something spiritual or contemplative; it is the success in taking into account all elements of the situation one is in:

Immersion wants to preserve only what is ‘pure’, essential, and enduring, while stripping away everything else; the concentration of which I speak does not consider our instincts as too impure, the sensuous as too peripheral, or our emotions as too fleeting – everything must be included and integrated. What is wanted is not the abstracted self but the whole, undiminished man (Buber 1970, p. 137).

The capacity to be in an I-Thou relation is not something that must be studied and practiced but rather something that must be re-discovered. It is approached not by acquiring new capabilities and skills but by attempting to break the armour that one gradually builds (Buber 1930b, p. 12) and by allowing the signs with which the situation discloses itself. Putnam’s misunderstanding of Buber is understandable since Buber repeatedly remarks that it is difficult if not impossible to answer the demands of the situation in a completely adequate way. Buber’s human being is always a fallible, searching, insecure agent. This should not lead to a further misunderstanding however, that what Buber points at is some paradise lost. The fact

that Dialogue and I-Thou relationships decrease is both a phylogenetical normality and part of Buber's *Kulturkritik* regarding the human being in modernity "Our age is intent on escaping from the demanding 'ever anew' of such an obligation of responsibility by a flight into a protective 'once-for-all'" (Buber 1936, p. 247 / Buber 2002, p. 82). He is anything but sentimentally bound to the idea that there could be a return to a period before the gains and losses of modernity: "We can as little return to the state of affairs that existed before there were schools as to that which existed before, say, technical science" (Buber 1926, p. 106).

Part of Buber's critique of his age, i.e., of modernity, is related to the increase in the tendency of humans to create "hideouts" for themselves. Faced with the demands of the situation, which, in Buber's description of the Chassidic way of life is understood as a call from God, the agent may fail to respond taking refuge in such a "system of hideouts" (Buber 1996, p. 11). But this is not simply an individual failing. Buber's interpretation of the history of ideas follows a similar trajectory and involves a number of attempts at establishing secure hideouts. Buber regards Aristotle's cosmology and Aquinas' theology as examples of such systems. Hegel's is the third great attempt at building such a "house": "All insecurity, all unrest about meaning, all terror at decision, all abysmal problematic is eliminated" (Buber 1938, p. 166). Buber's moral philosophy is critical of such a yearning for a safe haven and expects the moral agent to be on the move.

It might be true that Buber's ethic is demanding in this sense but in another sense it is not more demanding than can be expected of human moral agents. Human beings are fallible and it is practically impossible to answer the call of a situation in a morally immaculate way. What Buber asks of moral agents is nothing more than "not [to] do more wrong, than we must, to live" (quoted in Simon 1963, p. 502,

translation B.S.). Friedman paraphrases this principle as “*quantum satis*”; nothing is asked of the moral agent, which goes beyond her capacity and skill (Friedman 1963, p. 169). He quotes Buber as once again selecting the agent herself as the judge of how much of the infinite demand is realistically achievable: “The measure and limit of what can be achieved in a desired direction cannot be understood but by going in this direction. The powers of the soul can be measured only in being used” (quoted in Friedman 1963, p. 169, translation B.S.). Although Buber believes that moral values have absolute binding power, he is not absolutistic in what he expects of the moral agent and demonstrates a comparable attitude also when it comes to more abstract, collective entities and their social and political decision-making processes. His remarks on “his socialism” can be applied to his philosophy in general: He is “not a perfectionist but a meliorist [...] what is decisive is what shall be and remain the direction of the always renewed melioration, ever adapting itself to the new historical conditions” (Buber 1964, p. 76).

An Evaluation of Buber

An evaluation of Buber’s legacy and its value for an enquiry concerning the implications, which a relational model of self such as his can have for moral philosophy is complicated by the fact that his body of work is broad and at times contradictory. It is additionally made difficult by the variety of aspects, which he brings into the debate and his lack of a clear system. Nevertheless, as I have endeavoured to demonstrate in the preceding chapter, there are a number of key elements in his description of the human agent and in his ethical stance, which I have argued can be used to interpret Buber more systematically than he claims is possible.

On the whole his rejection of a systematic methodology and apart from some interviews, in which he is explicitly asked to clarify contentious points, his essayistic and poetic style makes his writings open to misunderstandings. Some of his proclamations are so strange that they threaten to annul the value of the rest of his work. Although his attitude towards God and religion is very critical and it seems possible to interpret him secularly without doing injustice to his intentions, some passages are disconcerting: “Immer ist das Spenden auf der Seite des Religiösen, das Empfangen auf der des Ethischen” (Buber 1953a, p. 578). Many of the criticism that he responds to with some signs of indignation and protest that he could have been misunderstood are more legitimate than he cares to admit.

Methodological Concerns

The first point of critique is therefore methodological. Buber explicitly states that much of what he wants to convey cannot be expressed verbally or systematically. Instead, he appeals to the experience of his readers, attempts a dialogue with them and uses metaphors and stories. This is especially the case when the subject matter is the self or the *Wesen*. When discussing dialogue as a method of psychotherapy, for instance, Buber writes “A specific very important method of healing – existential healing – comes about through [encounter]: Healing, in which not only a specific part of the patient is grasped but the roots of his being” (quoted in Reichert 1996, p. 71, translation, B.S.). Should this way of approaching the nature of the self through metaphors be understood as a necessity because it can’t be approached otherwise or is it carelessness on Buber’s part?

It is a general question, how that, which cannot be articulated or conceptualised should be treated philosophically. Wittgenstein's famous dictum suggests that we should remain silent. Buber on the other hand, tries to demonstrate by examples (Buber 1930, p. 145) or by appealing to the experience of the reader (Buber 1963, p. 593), when trying to show what cannot be formulated in concepts (Buber 1930, p. 145). This attitude is part of his insistence that he does not have a system or a doctrine (Buber 1963, p. 592).

The lack of system or use of a strictly defined terminology also makes difficulties when trying to understand the realm of the I-You. Since this realm is per definition outside the I-It, which is responsible for experience, explanation and description, the I-You can only be pointed at by referring to the readers own experience: "Since the perfect I-Thou relation in general makes no statement concerning itself, I do not know how frequent or how rare it is. I am concerned that the I-Thou relation be realized where it can be realized, and I cannot declare where it cannot be realized" (Buber 1964, p. 38). As such, the concept of I-You can easily fall prey to Bertrand Russell's tea-pot analogy (Russell 1952), as is the case more generally for the totality of Buber's (moral) philosophy.

Unfortunately, there is no satisfactory answer from Buber himself to such criticisms. Asked by E. La B. Cherbonnier if his philosophy is open to criticism, since "[t]he hallmark of philosophical discourse, as distinct from bare assertion or arbitrary insistence, is corrigibility" (Buber 1964, p. 50), Buber replies by saying that he doesn't reject consistency and that he is always open to criticism, without however replying to the more general accusation of being non-falsifiable.

A positive aspect of Buber's approach is his insistence on reflexive anthropology. Against the behaviourist reduction or elimination of the first person

view, which had a strong, if not predominant position in the first half of the 20th century, Buber argued consistently for both first personal and third personal views in a science of the human being. His own philosophy could be understood as adding a third dimension, that of the second person by the centrality of the I-You relation. One of Buber's merits is his procedural point that all of these dimensions must be included in philosophical anthropology and by extension in moral philosophy.

Another question concerning Buber's moral philosophy is the question of what exactly he is discussing. Lacking a clear system, his ethics make it difficult to distinguish between a variety of factors that are morally relevant. Marvin Fox aptly describes Buber's moral philosophy as an "analysis of the phenomenology of moral decision-making" (Fox 1963, p. 135). Yet instead of being a comprehensive phenomenology of moral decision-making it discusses one type, without it being clear, whether this is the only possible type or whether it is investigated as the normative ideal.

One of the advantages of Buber's moral philosophy is that it is melioristic and procedural. Although he believes that moral facts exist independently and that they have an absolute binding power, talk of morality in such a degree of abstraction is worthless for him. With the use of the term dialogical, he aims to point to a direction that human agents should aspire to but the direction does not involve principles or maxims. His comments rather postulate how the agents should be – attentive, open-minded etc. – and is thus related to the virtue ethicist approach without being identical with it.

Buber's description of a successful moral agent, his disinclination towards building a system and prescribing principles results in what I want to call a procedural moral theory. By "procedural" I mean that the emphasis of the moral

theory is not on what type of attitude or motivation the moral agent should have, nor on the possible consequences of an action nor on action-guiding principles or commands but rather on how the constellation of the moral situation and the moral agent had better be in order for it to result in a state of affairs that is morally better rather than worse. Such a theory does not postulate what is good or bad, right or wrong but is akin to virtue theory in emphasising qualities that the moral agent must have to be able to deal with the demands of the situation as well as is humanly possible. Buber's moral theory differs from other comparable virtue moralist theories because here moral situation and the moral agent are inextricably linked. His moral philosophy is procedural in that he argues how the moral agent should proceed in a given situation and facing a particular other. Thus his moral philosophy is concerned with the *how* of moral agency rather than with the *what* of moral principles or consequences. In addition to this procedurality, his absolutism with regard to moral values coupled with his epistemological scepticism makes his moral philosophy meliorist rather than overly demanding perfectionist: "I never thought an ideal dialogical relationship possible in our world as it is. I am a meliorist and not an idealist, and so I want only as much dialogic element as can be realized in human life here and now." (Buber 1964, p. 79).

Another formulation of the critique that Buber's moral philosophy is too demanding interprets him to mean that in every moral situation the agent faces a task of great searching and writhing. Buber denies that this is the norm however and his response to this point is similar to Mill's response to the criticism that utilitarianism would be infeasible because it would be too time-consuming (Mill 1987, p. 295):

One can catch the situation also in a flash, like the good tennis player can catch the ball, and like him can perform the correct countermovement in the same flash. I know and love some such people, in which everything has practically been decided

and the primal decision transforms into the actual and absolutely special as if from beyond the dimension of time (Buber 1963, p. 616, translation B.S.)

Does Buber's understanding of what the moral self is and what morality involves relieve us from approaching moral questions in a more conventional manner? Does his brand of particularism and situationalism imply that it is a vain undertaking to consider hypothetical moral cases, criteria for moral action, elements of the moral character or principles of moral education? I suggest that Buber's challenges to moral philosophy should not be understood as attempting to displace more traditional approaches but to add aspects to the discussion, which have been neglected hitherto. This becomes clear, when taking into account Buber's conceding, sometimes even advocating the use of norms and principles of the community one is born into. What Buber would object to in the use of principles, discussion around hypothetical cases and such is forgetfulness about the very fact that these are abstractions and that a moral theory based on such abstractions is worth little in the actual situation and its demands.

Key Concepts

Responsibility as a central term of Buber's moral theory fails to give substantial answers to concrete questions about how to act in a given situation. Thought experiments like the Trolley cases can be taken as a litmus test of whether a moral theory is able to respond to concrete questions on how to act. Buber's insistence that every situation is such a unique and complex constellation of factors that it is impossible to postulate principles or maxims, which can provide a matrix of possible answers involves the danger of being an empty truism. If the capability of responding to a situation does not involve either consulting general or even universal principles

nor reflexivity but rather an “openness” to the demands of the situation, then it becomes questionable whether the theory says anything substantial as a moral theory.

Buber’s concept of affirmation on the other hand involves two important insights. One is the postulation that the actual constellation of the self is not given but rather the product of a number of factors; mainly the individual’s personal biography, the situation and the quality of the encounter. It needs affirmation to be realised. The other postulation concerns the potential, “hidden” constellations of the self, which likewise need affirmation by other human beings in order to come to being and be able to do justice to the situation.

The central moral prescription in this context is “To do justice to (the reality that opens up to us in an unbiased manner)”⁶² (Buber 1995, p. 120, translation B.S.). For Buber, a situation, a encounter is done justice to, if the alterity of the other is affirmed (Buber 1950, p. 421). The individual is dependent on this type of affirmation to be able to be and to act. In turn, she must be willing to and capable of affirming the other, despite all possible attempts at influencing them. I want to argue that although this description of an encounter as genuine and necessary is accurate, it does not exhaust all types of genuine encounters. The very forgetting of the other in the immersion in a common project or a common goal can also be a case of affirmation. Paradoxically, I believe, affirmation of the other can reach its highest level, when it vanishes and both or all parties are turned to a common, third entity (an object, person or event). The manifest affirmation that is displayed here is one that does not focus on the 2nd person other but commonly on a 3rd person other and thereby indicates the existence of genuine affirmation between the primary parties.

⁶² As in the phrase “*der sich uns eröffnenden Wirklichkeit unbefangen gerecht zu werden*” (Buber 1995, p. 120)

Another point about affirmation is concerned with two legitimate approaches to the moral situation. When faced with a situation, in which one is to act, there are two aspects that are morally relevant. As Buber rightly argues, one concerns the other and her affirmation. There is another aspect however that may be but is not necessarily identical with the affirmation of the other. This is especially explicit in cases, in which one faces more than one other. In such a moral situation there is, beside the will to affirm the other, also the vision of how I want the moral situation to develop. There is another, potential moral situation latent in the actual one, which I want to affirm and which I want to be realised, while there are others I want to remain unrealised. In other words, Buber's recommendation of the affirmation of the potential constellations of self of the other can be applied to the moral situation as well. I argue that the vision, which of the potentialities of the present, actual moral situation should be realised, is as important an aspect as the affirmation of the other, who may or may not be interpreted as part of the moral situation.

Some of the main characteristics of Buber's views on the nature of self and on moral philosophy can be summarised around the key words of relationality and particularism. A first point of enquiry is the question why he does not bite the bullet and say that therefore the self of the human being at any of these given situations is a particular one. Why is this uniqueness of every situation and the being of the I as part of the pair that is unique to that situation not a reason to think that there is no such wholeness and unity of the self as Buber likes to claim? Is his scholarly resistance to be explained as part of that more common and unquestioned resistance towards the idea of disunity and dissonance, as portrayed in the many methods of self-building and self-protection, both pathological and non-pathological? There are indications that this might be the case, for he postulates that the difference between resolutely

being on the chosen 'path' and the immoral or amoral state of being is an unfocused hither and thither (Reichert 1996, p. 214). His insistence on the value of the 'path' might be an indicator that the unfocused hither and thither is to be shunned due to two reasons: It is an unpleasant state for the agent and it hinders her from having the focus and concentration necessary to respond adequately to the (moral) situation.

In Buber's view there is a direct link between the nature of the self and moral agency. The moral agent who finds herself in a particular moral situation will be able to respond the better, the more unified her self is, Buber claims (Buber 1947; quoted in Reichert 1996, p. 176). The more often such harmony is achieved, the higher the probability that it will be achieved again until one can safely rely on the unified self to respond appropriately. In short a life-long moral development might contribute to ever stronger unity and correspondingly to ever better responses. A mark of this development is that intrapersonal conflicts and contradictions will be overcome (ibid.). What precisely does it mean however, that conflicts and contradictions will be overcome and what is the process behind such an event of overcoming? Should the conflict be understood as one between two parts of the self and if yes, does this mean that Buber proposes a model of the soul that is similar to the Platonic one? How then is the conflict mastered? Does one part of the conflicting parties give in and vanish? Unfortunately Buber does not provide answers to more detailed questions about how the unity of the self, necessary for (moral) agency is achieved and maintained.

One of line of Buber's thought, which can be understood as an argument for this postulation of unity is his conception of the "real self". Much of his moral and religious philosophy is based on his conviction that the self can be more or less authentic in the sense of being more or less alike to what its 'real nature' is.

However, the “real self” in Buber is different than other, more traditional postulations of self, because it is only “real” when it is in a relation; „the real self only appears by forming relations to the other and where this relation is revoked, the self dies“ (Buber 1953a, p. 577, translation, B.S.). Much, which has a religious tint in Buber’s writings can thus be secularised by referring to the concept of authenticity. When he elaborates on what he means by “holy” or by “hallowing”, for instance these appear to be approaches to the other without preconceptions, spontaneously, with more focus on the other, i.e., with more attentiveness and with more open-mindedness (Buber 1964, p. 62). Whether the postulation of a unified self is thus answered by the postulation of a real self remains an open question however. If the premise that the relational self can be more or less authentic in a given situation is true, it remains unclear why this authenticity should involve unity.

More specifically such questions concern Buber’s description of the *Wesen*, which acts best only when it is involved in the situation, in an I-You relation, with the whole of its *Wesen*. What happens with the rest of the *Wesen*, when she is in an I-It relation? Had this rest better be understood as nothing or as non-*Wesen*? Does the rest consist of unactualised, potential parts of the self? Buber’s view can be little more than speculated upon, since he does not take up this question explicitly and only in passing: “The self-knowledge of the one who turns towards [something] of his turning towards, this recourse of the Rest-I, which does not take part in the act, of which it is an object, de-possesses the moment, de-spontaneities it. [...] One has to understand this correctly: this is not simply about a special case of the well-known sickness of the modern human, to have to take part at the own actions as a spectator” (Buber 1953a, p. 597). Buber’s diagnosis is rather that of a sickly loss of faith and trust in the absolute.

While discussing Buber's frequent praises of wholeness, unity and the like, there is some danger of misunderstanding his meaning. For example, he also uses the word "wholeness" (*Ganzheit*) when discussing the aim of philosophical anthropology and criticises Kant for disregarding it (Buber 1948, p. 311). What he means here does not seem to be so much the unity of the different aspects of human life but rather the comprehensiveness of its study. In some passages at least, what Buber means by the word 'wholeness' seems to be a quality of the philosophical anthropological approach to the human being, which should take into account the totality of the characteristics that make up the human being. The present practice tends to get to work by an attempt to philosophically divide and conquer; "Either a philosophical discipline shuts out man in his complex wholeness and considers him only as a bit of nature, as cosmology does; or (as all the other disciplines do) it tears off its own special sphere from the wholeness of man, delimits it from the other spheres, establishes its own basic principles and develops its own methods" (Buber 1948, p. 145). In this sense, what Buber criticises as lack of wholeness is not some sort of inner coherence of the human being as a mind-body complex but the specialised focus on one or some aspects of human life.

Buber's Moral Theory

Two key concepts, which can be used to describe Buber's thought; i.e., situationism and particularism are directly related to Buber's moral philosophy. These two notions lead him to dismiss moral systems not only as useless but as positively dangerous. Morality as a set of ready-made prescriptions about how to live and how to act is the main obstacle for a moral agent's perception of the other and of the

situation (Buber 1930, p. 165). Both morals and religion deny the moral agent the chance to experience the situation as unique, by teaching her and habitualising her in patterns of perception and evaluation “Principle there, dogma here, I appreciate the ‘objective’ compactness of dogma, but behind both there lies in wait the – profane or holy – war against the situation’s power of dialogue, there lies in wait the ‘once-for-all’ which resists the unforeseeable moment” (Buber 1930, p. 21). The addressee of this warning is not only traditional or religious moral systems but also academic occupation with morality if academic ethics disregards the realm of the I-You and sticks to that of the I-It.

Buber’s own answer to one fundamental question of ethics is therefore characteristically holistic or comprehensive. When a moral agent finds herself in a situation, the primary decision that she has to make and generally makes automatically and unconsciously is whether the situation is a moral situation, whether there is a moral demand with which she is faced. While it is possible to consider this ethical demand as either moral in the narrow sense as affecting the well-being, interests or rights of others or as a question of the aesthetics of life, Buber disregards this differentiation. He pleads for as much care and consideration as possible in his relations firstly with the what Buber calls “the world and things”, secondly with other human beings and thirdly with the Absolute/God:

In virtue of his nature and his situation man has a threefold living relation. He can bring his nature and situation to full reality in his life if all his living relations become essential [*wesentlich*]. And he can let all elements of his nature and situation remain in unreality by letting only single living relations become essential, while considering and treating the others as unessential (Buber 1948, p. 210).

Buber’s preoccupation with God or the absolute sometimes evokes the idea of a special realm of life, in which encounters with others take place but this would be a misunderstanding. He emphasises a number of times that it is the everyday life,

which is the focus of his investigation, not the exceptional situation and similarly the totality of the others with which one becomes the I of an I-It or an I-You couple is simply “the medley of the nameless crowd” (Buber 1948, p. 211). In everyday life, the trick to answer the moral demand is to refrain from singling out some relationships as primary while functionalising or objectifying others; i.e., by categorising some as inherently and necessarily belonging to the realm of the I-It: “When we do not resolutely effect the distinction between the two attitudes, we further, even if very much against our wills, the tendency which has grown so strong in our time to ‘manipulate’ the existing being” (Buber 1964, p. 21).

A further merit of Buber’s moral philosophy is his constant reminder that the encounter of the I with the You is not the product of a ‘higher’, more developed state of either individual or society. Therefore he approaches the traditional mind-body dualism, that appears in a variety of guises, as in Schopenhauer’s understanding of the world as will and idea, or Scheler’s adaptation as impulse and spirit with critical distance:

The true negotiations and decisions take place, in the life of these and in general of great men, not between spirit and instincts but between spirit and spirit, between instincts and instincts, between one product of spirit and instinct and another product of spirit and instinct. The drama of a great life cannot be reduced to the duality of spirit and instinct (Buber 1948, p. 226).

In moral practice this means that the child who approaches an animal as a You will succeed in answering the demand of the situation, without necessarily having an account of moral principles or habitualised virtues.

These aspects of Buber’s moral theory have the advantage of localising morality more broadly, i.e., more situations are considered as morally relevant, more individuals are considered as possible moral agents and more entities are considered as morally relevant others or moral patients. This becomes clear in a comparison of

Kant's moral position and Buber's. Buber's insistence on giving the other the primary focus, his repeated pleas not to allow the realm of the I-It unchecked growth to the detriment of the realm of the I-You, his warnings of the dangers of the objectification, his pleas for one's own authenticity concentrate around a very genuine respect for the authenticity and freedom of the other, "I live 'ethically' when I confirm and further my Thou in the right of his existence and the goal of his becoming, in all his otherness" (Buber 1964, p. 28).

The similarity between one version of the categorical imperative – "Act in such a way that you treat humanity, whether in your own person or in the person of any other, always at the same time as an end and never merely as a means to an end" – and Buber's own position – "not to treat the other as an object but rather as his partner in an instance of life, even if it is simply in a box match. This is the key point: the not-being-an-object" (Buber 1954, p. 274) – led Buber to comment on the difference between his championing of three elements of the realm between two human beings; i.e., firstly the lack of appearance, secondly the affirmation of the other and the will that she may realise what is proper to herself and thirdly abstaining from imposing on the other as in the example of propaganda (Buber 1954, p. 291). Although this much seems to accord with Kant's position, Buber emphasises that his central point is concerned with the preconditions of what happens between human beings, "The human being does not exist anthropologically in his isolation but in the totality of the relationship between one and the other: only the interaction makes it possible to grasp humanity adequately" (Buber 1954, p. 290, translation B.S.). This goes beyond localising the moral element in the will of the agent and encompasses the moral agent, the other – who may or may not be considered as a moral patient – and the situation.

Buber's description of the moral situation and the role of the moral agent in it as, in a way, quasi phenomenological analysis has an important advantage over more traditional moral theories. Both deontology and consequentialism provide the moral agent with a much greater sense of security than it's possible to have. Some virtue ethicist accounts of morality, while allowing the situation and the particular position of the agent in the situation more importance, nevertheless have the tendency to put practice and habitualisation in the centre of morality. The picture that Buber draws of the moral agent does justice to the lack of direction, courage and focus that may befall the moral agent. At times the conflicting parts of the self may bond together and the moral agent may act in a more unified way but even this intra-individual co-operation does not provide any security for the fundamentally alone agent, there is no "universally valid answer, no sort of guarantee, only a chance, only a venture" (Buber 1963, p. 618, translation B.S.).

However these finer details had best be interpreted, a great merit of Buber's moral philosophy is his analysis of the moral situation as a compound of agent, patient and situation and his localisation of this compound in the everyday life. Connected with this general stance is his critique of inauthenticity and its effect on morality. His analysis of inauthentic relations pertains not only to the will of domineering over the other but also to the enjoyment of the other. Such enjoyment is not only typical of erotic relationships but also a danger in relationships between educator and educand (Buber 1926, p. 113). The danger involved in such relationships is that of a "falsification, beside which all quackery appears peripheral" (ibid.). Translated into daily action, this means that the authentic response to a specific other cannot be grasped with general maxims as for example, to turn the other cheek, be understanding, tolerant or kind.

The Position of the Moral Patient

A way to test Buber's moral theory is to consider the position of the moral patient in it. And although the other is accorded such a central role in Buber's thought, although it's the other that is the starting point of the moral situation rather than the agent, it would be a mistake to identify the other with the moral patient. Buber's particular ontology, which rejects the subject-object dichotomy would also reject a description of the moral situation that includes a denomination of its cast as clearly separable moral agents and patients. What both – or all – parties in a given situation should be is as authentic as possible to allow the other to become more authentic. This means that one has to provide help to the other to enable her to realise herself more thoroughly. In a passage from an early essay, Buber generalises the help that is needed by everyone into the primary demand of the other

You shall help. Each man you meet needs help, each needs your help. That is the thousandfold happening of each moment, that the need of help and the capacity to help make way for one another so that each not only does not know about the other but does not know even about himself. It is the nature of man to leave equally unnoticed the innermost need and the innermost gift of his own soul, although at times, too, a deep hour reminds him of them. You shall awaken in the other the need of help, in yourself the capacity to help (Buber 1919, p. 110).

Is it conceptually legitimate then, to identify the other as the moral patient, even if only for arguments sake? If it is, is the Buberian approach to the other a satisfactory way of treating the moral patient? I believe it would not be a grave misreading of Buber's thought to interpret some of the entities that function as the other as a moral patient. Although Buber's ontology should make it impossible to speak of an I by itself, there are enough indicators to argue that the other is a creature for whom things can go better or worse, who may be in need of help, who may be treated in a way that is better or worse for the realisation of what is proper to them. This makes it

possible to treat the Buberian other as a moral patient for our purposes. Not every “other” in Buberian thought is a moral patient however. While we can argue that every moral patient can be regarded as an ‘other’ in Buber, the opposite is not the case. The category of the other is much wider and encompasses non-sentient beings like trees, collective, abstract entities like nature or non-living beings like art.

A further question regarding the position of the moral patient in Buber’s moral philosophy concerns his postulation that the person who has been in an I-You relationship comes out of the experience as someone who is in a measure transformed. I-You relationships leave their mark on the I-It realm and influence the (moral) judgement and behaviour that follow. If one regards the treatment of a particular moral patient who has been in an I-You relation with a moral agent, is their evaluation as an object in the I-It realm changed as well? Does the agent value the other more if she has related to it as a You at some point in the past? A challenging case would be to consider a child and a toy or a child and an animal. These are among the partners with which children tend to form I-You relationships but it is just as common for children to discard them afterwards as to cherish them more. In other words, does the fact that a moral patient has been the other, the Thou for a moral patient, ensure its better treatment afterwards? Buber claims indicate that this ought to be the case but whether it is empirically true remains very open to doubt.

A central element about encounters with a variety of others that might affect the moral position of the other as a moral patient is mutuality and reciprocity. Buber considers reciprocity to be important enough to equate it with relation: “Relation is reciprocity” (Buber 1970, p. 67). He thinks that there “are several different grades of the capacity for mutuality” (Buber 1964, p. 37); i.e., although a sculpture, a chicken or a philosophy professor can all be the other in a given encounter, the degree of

mutuality, of reciprocity, of the responsiveness of the other are not necessarily all the same. Nor is the degree of mutuality a sufficient ground to determine the response to the other and to the situation, “[...] I am by no means of the opinion that from this alone a ‘basis of evaluation’ can be established. To this end, rather, our whole knowledge about the world must co-operate, a knowledge that is ever again renewed through the I-Thou relation, but is not borne by it” (Buber 1964, p. 37).

Mutuality can have degrees then and it is not necessary, not even possible that it should be “complete” (Buber 1970, p. 179). Nevertheless, mutuality remains central:

[...] even if speech and communication may be dispensed with, the life of dialogue seems, from what we may perceive, to have inextricably joined to it as its minimum constitution one thing, the mutuality of the inner action. Two men bound together in dialogue must obviously be turned to one another, they must therefore – no matter with what measure of activity or indeed of consciousness of activity – have turned to one another (Buber 1930b, p. 9).

This point is one of the more problematical points of Buber’s thoughts and remains obscure. For the other, as Buber remarks in a number of places, can be a work of art or an object in nature as much as another human being. Since “turning to one another” cannot be expected to take place between a human being and a mountain or more problematically between two mountains, what does this tension between the possibility or the degree of mutuality involve for the position of the patient?

How then does this moral patient, this other fare in a world of Buberian moral philosophy? If we subscribe to Buber’s anthropology, ontology and ethics, what is the position of the moral patient in it and is it an advantageous position? Buber’s rejection of universally valid values in his moral philosophy demands that the attitude towards moral agents may not be one of imposition (Buber 1953, p. 289). This dictum is part of Buber’s conviction that “in every person, the right thing is invested in a singular and unique personal way” (Buber 1953, p. 289, translation

B.S.). Once again, this conviction of Buber opens the question whether the treatment of moral patients does not become arbitrary. If one's primary concern is with the well-being of, for example a cow, then how can one refrain from imposing or wanting to impose attitudes and behavioural patterns on others who might not share one's concern for the cows well-being? Of the three ways of influence on other moral agents, education remains the only one that is open in a Buberian framework, in order to pursue my aim of increasing the well-being of moral patients.

This is a problem for Buber's moral philosophy. It is also related to the relationship between care ethics and Buberian ethics. One can only speculate how Buber would react to the challenge. Buber, to recapitulate, argues against the Heideggerian notion of care (*Fürsorge*). Levinas agrees that Heidegger is not the person to turn to when one wants to learn about love for human beings and social justice (Levinas 1963, p. 131), although this is probably a comment on Heidegger's affiliation with the Nazis rather than a philosophical point. Nevertheless, "care as a response to an essential affliction is an access to the alterity of the other. [...] One may ask if to clothe nakedness and satisfy hunger isn't the real concrete access to the alterity of the other – indeed, more real than the ethereal air of friendship. Is dialogue possible without care?" (Levinas 1963, p. 131). This question remains unanswered as does the question what Buber would consider to be the best response for a moral agent who finds herself faced with another human being about to attack and kill a cow.

I suggest that it is a mark of a good theory if it can provide a good answer to the problem of what I shall call the moral brute. This is the general question concerning when and how one ought to or might be allowed to interfere in the actions of others and with what legitimisation. It is related to but not identical with

another challenge to moral philosophy, the amoralist, who personifies one of the grand questions of moral philosophy “Why be moral?”. The only indications about Buber’s position can be found in an interpretation of certain passages as dwelling on the problem of the moral brute. It would be a misreading to interpret him as advocating a comfortably liberal position of non-interference in cases where a moral agent can be said to act like a brute. On the contrary, he emphasises repeatedly that taking the other seriously and being in a dialogue with her will sometimes involve fighting the position of the other. What is being affirmed is a set of possibilities that is not actualised in the present demeanour of the other. Is this a satisfactory answer? If it is not the actualised position of the other that I affirm but the hidden potentialities of her wider self, what makes the latter worthy of being affirmed? Can I really abstain from completely and utterly rejecting the brute by separating her actual brutality from her further potential self? Furthermore; on what grounds do I pass such judgement? Buber’s dictum is similar to the Abrahamic dictum “Hate the sin, love the sinner” but this is not possible unless I provide reasons to support which aspect of the others self should be affirmed and which rejected.⁶³

Martin Buber’s resistance to systems, principles or theories is mirrored in his rejection of normativity with regard to the dialogical life. In a passage, in which he answers a fictional critic Buber emphasises that he does not and cannot put demands. Responding to or answering to the situation is not something that can be demanded or furnished with an ‘ought’ (Buber 1930, p. 190). Human beings are able of being attentive and open-minded, Buber claims, without these characteristics being in any way connected to so-called higher faculties or experience. The agent does not even know herself with regard to how well equipped she is to perform these feats

⁶³ The converse is also true; there are no “Good” people on earth, there is only the Good (Simon 1963, 500)

necessary for moral action and can – possibly surprisingly – only realise this in the situation that puts a specific demand to her (ibid.).

Buber's position about the brute can also be guessed at from his remarks on Hitler. The other can be the antagonist with whom fighting is a way of affirmation but not every (human) being is automatically an other. Buber's description of Hitler is an example for this:

Hitler is not my antagonist in the sense of a partner whom 'I can affirm by contradicting him', as Friedman says, for he is incapable of genuinely addressing one and incapable of genuinely listening to one. I have experienced this one personally, when I heard him speak, be it through the technical medium of the radio. I knew that this voice was capable of destroying me together with countless of my brothers; but I experienced that despite such power it was not capable of putting the spoken and heard word out into the world. Und barely half an hour afterwards have I begun sensing the 'poor devil' in the 'satan', the poor devil at power and at the same time I understood my dialogical helplessness. I had to answer but not to the one who had spoken. Insofar a person is part of a situation I have to answer but not specifically to the person" (Buber 1963, p. 622, translation, B.S.).

If we return to the example with the moral and the cow, this position indicates that the brute can be considered as a "poor devil", who loses the ontological position of being the other whom I can oppose or affirm. To use another terminology, she can be considered as neither moral agent nor moral patient. This unsatisfactory state might indicate the need for a third category for moral entities like moral tornados.

The Place of Principles and Consistency in Buber's Moral Philosophy

Buber's moral philosophy is open to the objection that following Buber's call to be open to dialogue, to be attentive to the demands of the situation and to answer in a way that is responsible towards the other could generate inconsistency as well as arbitrariness. Since he is sceptical about the worth of absolute principles in the moment of a concrete moral situation that demands action, the variety of choices that

can accumulate over a lifetime may be far from having a common denominator, let alone a constant and consistent moral agent behind them.

As it stands, this criticism is too vague and allows the extraction of two distinct objections. One is the allegation that there is a lack of practical consistency in the totality of actions of a Buberian moral agent. To this objection Buber responds by indicating the unique singularity of every situation. The criticism of practical inconsistency would imply that situations and actions are similar enough to be compared. This assertion however is precisely what Buber rejects. No two situations, no two agents and no two actions are similar enough to be compared and it is impossible, even theoretically, to determine principles around which to prescribe the right moral action. In an interview where he is questioned about his lack of principles, Buber remains firm “I have no principles. There are things that we must do – here and now. I have no principles, only a sense of orientation and I act according to the given situation. [...] A principle is something that forces you to always act in a prescribed way [...] every situation must be observed anew. [...] Keep your eyes open – that is all that I have to tell you, for I cannot name you a principle” (quoted in Reichert 1996, p. 88, translation B.S.). To the insistent question considering the criteria on which to base judgement Buber repeats “You have to observe the situation. I cannot make it easier for you” (quoted in Reichert 1996, p. 89). And:

“In spite of all similarities every living situation has, like a newborn child, a new face, that has never been before and will never come again. It demands of you a reaction which cannot be prepared beforehand. It demands nothing of what is past. It demands presence, responsibility; it demands you.” (Buber 1939a, p. 135)

The obstinacy that Buber demonstrates in this interview includes his persuasion that his view of morality is the more difficult path. While cases of moral failure around

theories based on principles concentrate on cases of temptation to ignore principles and follow inclination, Buber's understanding of *akrasia* is more situational. It is easier to determine a set of principles and follow them, easier to live a life of practical consistency than of inconsistency. As tempting as it may be to succumb to inclination, to make an exception, to convince oneself that the principle doesn't apply to a particular case, it is much more difficult to face the existential loneliness and constant awareness that the Buberian moral agent has to face.

The charge of inconsistency is therefore misapplied when directed towards a hypothetical Buberian agent but it might be less so when directed towards Buber himself. Throughout his long life Buber has produced a vast bibliography on a variety of subjects ranging from philosophy to theology, from politics to Judaism. His declaration that he does not have a system might lead to a challenge of inconsistency about some of his dictums. Despite his proclamation that he cannot prescribe anything, for example, sentences like "You shall help" (Buber 1919, p. 110) or "You shall not withhold yourself" (Buber 1919, p. 109) are clear prescriptions. His call for situationism can be interpreted as a principle itself (Reichert 1996, p. 88) and although he insists that the correct way to respond to a situation is to be open-minded, attentive and responsible towards the other, he does not only not condemn acting according to norms and principles. One answer that can be given to such a criticism of inconsistency is the very length and breadth of Buber's work. To expect comparable positions from Buber in his early twenties and from Buber in his late eighties would be stretching even his ideal of a unified and constant agent too far.

Another possible answer would be to question the realisability and worth of consistency itself. If Buber's argument against principles could be applied against

consistency in general, this would involve the complexity of life, the great number of factors that exist in every situation, the impossibility for an agent to take all these factors into account, the fallibility of even the most open-minded, most earnest, most anxious of Buberian agents, the impossibility of postulating an outside world as it is that can be recognised and understood. Would this reminder of the ontological complexity of the world and the epistemological inadequacy of the human agent be enough to abandon consistency as an ideal impossible to reach? An alternative might be to dispense with the claim to absoluteness and regard consistency as achievable in degrees and strive for the best.

It is an indisputable fact that there is a great variety among moral agents. This variety pertains not only to factors like preferred moral theory, degree of practical consistency in moral situations, techniques of moral judgement and the like but to the very readiness to behave morally or the willingness to regard morality as a legitimate and necessary component of agency. Similarly, if we apply Buber's terminology, there is a great variety in agents' readiness and capability of entering into dialogue. When asked by Perry LeFevre, what factors are in play in creating this variety among people, Buber's answer is far from satisfactory. He states that "[t]his is a field into which I can venture only with difficulty" (Buber 1964, . 29) and contents himself with describing some of the differences between two polar types; i.e., of agents who remain as independent and secure as possible on the one side and agents who take the risk of encountering others and allowing (inter-)dependency to develop. The question why there is such variety and how it can be explained, whether as a result of choice, education or innate factors remains unanswered. Although Buber's moral philosophy is based on his moral ontology and his conception of the human self, there is little of moral psychology as practiced today.

A questionable element of Buber's moral philosophy is his understanding of good and evil. To recapitulate, Buber argues that:

Evil is lack of direction and that which is done in it and out of it as the grasping, seizing, devouring, compelling, seducing, exploiting, humiliating, torturing and destroying of what offers itself. Good is direction and what is done in it; that which is done in it is done with the whole soul, so that in fact all the vigour and passion with which evil might have been done is included in it (Buber 1953, 130).

CHAPTER 5

ARNE NAESS AND THE ECOLOGICAL SELF

Arne Naess is the second example of a philosopher who argues from a relational self to a certain understanding of normative morality. He and Martin Buber have a great deal in common: Pluralism, interdependence and relationality could be used as keywords to both their philosophies. Apart from the content of their thought, their perception as philosophers is another characteristic that they can be argued to share for both are unusual philosophers and considered themselves as such. While Martin Buber considered himself to be an atypical philosopher, Arne Naess called himself a “‘philosophical vagabond’ [...] what the ancient Greeks called a zetetic” (Drengson 2008, p. 19). Although both had academic positions as philosophers, they were critical of the figure of the professional philosopher and were both well-known as activists outside academia. In fact, both were considered as somewhat unlikely heroes in their respective countries; Martin Buber was buried with a state funeral in Israel, which was moreover attended by members of the Arab Students Delagation while Arne Naess is described as having been regarded as a “national treasure” (Drengson 2008, p. 4). Their longevity together with their study of modernity and its problematic, if not to say disastrous legacy make both of them characteristically 20th century figures.

Pluralism and Critique of Modernity

Naess joins Buber in his spirited critique of modernity and its effects on human life but while Buber concentrates on interpersonal relationships, for Naess the ecological aspect of selfhood is more central. This informs one element of his critique of modernity as the loss of an identificatory relationship between self and a specific place. Through the gradual loss of traditional forms of living, Naess argues, humanity came to suffer from “a place-corrosive process”⁶⁴ (Naess 1992, p. 45). In an intact relationship with the place that one lives in and which one existentially depends on, the environment becomes literally a part of the self. In modern life the development of such an extended self is endangered:

Urbanization, centralization, increased mobility (although nomads have proven that not all sorts of moving around destroy the relation of belonging somewhere), the dependence on goods and technologies from where one does not belong, the increase of structural complication of life – all these factors weaken or disrupt the steady belongingness to a place or even hinder its formation (Naess 1992, p. 45).

Naess seems to echo Buber in his postulation that an extended self that is shaped by the demands of the situation – in this case the place – and his rejection of an extensive subject-object dualism: “Phenomenologically speaking, the orders given by the place and the orders give by oneself are inseparable. Only philosophies that impose a sharp subject-object dualism try to trace a border between the self and ‘its’ geographical surroundings” (Naess 1992, p. 57). When aboriginal people complain that a part of them is destroyed when their traditional homelands become areas of the development of dams, hydroelectrical santrals and the like, Naess argues that this

⁶⁴ Naess’ discussion of the importance of an intact relationship between a place and an individual for the formation naturally also takes into account the less alluring aspects: “It [a favoured place, B.S.] may enrich life, but may also lead to a manifold of habits and ways of thinking that are peculiar and a source of irritation to anybody not adapted to that special life. I find that attachment to places should not be praised uncritically” (Naess 1992, p. 60).

should be understood quite literally. The tendency to think that the individual and the place are linked in such a way as to make the formulation of the complaint a metaphor is grounded in a continuation of “thinking of two completely separable, real entities, a self and the place, joined by an external relation” (Naess 1986, p. 88).

Historically, Naess considers this process to have begun with the Renaissance, which “[...] glorified our ego by putting it in some kind of opposition to the rest of reality” (Naess 1986, p. 88). Nevertheless, it’s not only modernity that is the target of Naess’ critique but also the critics of modernity, Heidegger, Sartre, Kierkegaard, Wittgenstein, Marx and Nietzsche. Since Naess understands maturity as expressed in a sense of a wide, ecological self as opposed to the more narrow concept ego, these philosophers are guilty not only of immaturity but relatedly of a lack of integration (Naess 1973, p. 125). The way out of the alienation that has been among the main focus points of the critique of these philosophers, coupled with the existential Angst and *Weltschmerz* is “high degrees of integration of the personality, and high degrees of such integration require intense cultivation of the personal aspect of interaction with the environment” (Naess 1973, p. 127). Like Buber, Naess does not romanticise a better past and considers that a return might be possible. His goal is ecological sustainability and he approaches this goal pluralistically: “[...] there is not just one way but many ways, so that widely different, sustainable cultures are possible” (Naess 1995, p. 290)

Naess’ pluralism in envisioning an ecologically sustainable future is part of his pluralist approach in general: “According to Naess, there is never one definitive interpretation of philosophical texts; there is never one single description of an event or a single theory of things that is the whole and only truth” (Drengson 2008, p. 20) but while the accounts of it are plural, Naess nevertheless considers reality to be

“one” (Naess 2006, p. 182). In general, the link between pluralism, i.e., the idea that there are a plurality of perspectives or values to be taken into account when describing a specific state of affairs or prescribing another and between a normative stance that demands respect for the other or a non-violent, non-interfering approach seems common to both Buber and Naess – and incidentally to Mill. All three argue that it is not possible for a single human being to cognise truth, either because truth is not an independently existing realm of epistemically accessible facts or propositions or because it is multi-faceted with each individual only having access to one perspective at a given time. This view does not only regard the value of the other’s perspective or taking part in a given situation as intrinsically valuable but also as instrumental to the creation or increase of a good.

Ecosophy

Naess’ position on the nature of the self and its implications for morality are embedded in his wider philosophical position. Three terms are intimately connected with and generally used to denominate his philosophical stance: Ecosophy, deep ecology and ecology of wisdom. That these terms are interconnected becomes already obvious from the fact that “eco-“ is included in all three and ‘wisdom’ in two. The inclusion of wisdom is a shortcut for two main elements: Firstly it is “openly normative, it contains both norms, rules, postulates, value priority announcements and hypotheses concerning the states of affairs in our universe” (Drengson 2008, p. 32). Secondly Naess’ understanding of wisdom is comparable to the Aristotelian concept of *phronesis*; it involves the ability and the experience to

perceive a situation accurately and additionally the readiness and capability of making appropriate (axiological) judgements.

Ecosophies are individual and dynamic, i.e., changing philosophical systems⁶⁵ developed by individuals and denoting their philosophical stance both on how they interpret the world to be and how they consider it ought to be. Naess coins the word ecosophy by referring to the concept of wisdom as well as the term *oikos*, meaning household, house, community or family familiar also from the stoic concept of *oikeiosis*. Naess broadens the scope of the term and includes also the earth in its meaning: “So an ecosophy becomes a philosophical world-view or system inspired by the conditions of life in the ecosphere” (Drengson 2008, p. 32). Naess calls his individual ecosophy “Ecosophy-T”, where the ‘T’ alludes to Tvergastein, Naess’ spartanic mountain hut in the Hallingskarvet mountains in Norway, where Naess spent much of his time and which he considered as a component of his wider self.

Naess’ Conception of Self

Naess’ approach to the question of self is not only that of an inquisitive philosophical scholar but also that of a deep ecologist with an ethical agenda. His discussion of the self therefore involves an elaboration on the question of agency, i.e., on the question of what type of self promises to be associated with better or worse moral agency. Self-realisation is the supreme norm in his brand of ecosophy but Naess emphasises that what is meant by ‘self’ is not a substance with the property of permanence, as he criticises Descartes and Spinoza to have taken it to be (Naess 1985, p. 195). He himself prefers to understand the self as an “*ens rationis*”, i.e., as an “abstract

⁶⁵ Naess is decidedly not averse to establishing systems as Buber is.

construction [...] created (by reason) to facilitate rational analysis” (ibid.). Naess positions himself near to Gestalt ontology and to Buddhism and describes selves as “frequently recurring entities, or ‘knots’, in the structure of contents” (Naess 1985, p. 196).

Naess discusses one Buddhist way of considering selves as “processes or aspects of processes, always changing, but always showing an important, limited continuity and permanence” (Naess 1985, p. 197). A similar formulation can be found in Naess’ discussion of how the subject-object dualism is transcended in Buddhist thought, where “a happening refers to a whole constellation or gestalt of relations” (Naess 1985, p. 199). The Buddhist view of ‘not-self’ (anatman⁶⁶) clearly demonstrates one link between the conception of self that a moral agent has and their moral stance. Exemplary for this link is a quotation of Rāhula, a Sinhalese monk:

According to the teaching of the Buddha, the idea of self is an imaginary, false belief which has no corresponding reality, and it produces harmful thoughts of ‘me’ and ‘mine’, selfish desire, craving, attachment, hatred, ill-will, conceit, pride, egotism, and other defilements, impurities and problems. It is the source of all the troubles in the world from personal conflicts to wars between nations. In short, to this false view can be traced all the evil in the world (Collins 1982, p. 4).

Naess understands the self as relational and as consisting of plural components. He churns talk of “the plurality of selves” (Naess 1986, p. 89) and prefers to conceptualise the self as having multiple components. These components are relational; i.e., it is the relations that an individual has to “other people, to material things, and, certainly, to what we call his or her environment, the home, the garden, the neighborhood” (ibid..) that constitutes these components of the self. As with

⁶⁶ One view is that what ‘annattā’ or ‘anātman’ refers to is not the non-existence of a thinker of thoughts, experiencer of experiences and agent of actions, but rather a specific ego, which is different either from the body (Collins 1982, p. 7) or from the ego understood as a persisting, unified entity (Collins 1982, p. 9). Thus, the ātman is sometimes understood as “The divine element in Man, degraded into idea of an entity dwelling in the heart of each man, the thinker of his thoughts and doer of his deeds [...]” (Collins 1982, p. 8).

Buber, the 'others' of the relation are by no means other human beings; they can be animate or inanimate parts of the environment. This understanding of this sort of wide self also grounds Naess' epistemological position. The self is not the subject who has access to reality as a thing in itself but rather only as "separate parts or aspects in separate moods" (Naess 1986, p. 94). This does not mean that the self is fragmentary or that it is desirable that it should be so (Naess 1973, p. 123). On the contrary, Naess considers the specialisation and uneven maturation that modern life seems to force on individuals as something that needs to be overcome in order to avoid alienation as discussed by social critics of the enlightenment and the industrial revolution.

Naess' analysis that "We underestimate ourselves. And I emphasize selves. We tend to confuse our 'self' with the narrow ego" (Naess 1986, p. 81) leads him to widen the conception of self. To the more commonly postulated views on a wider self, i.e., a social self or a metaphysical self, he proposes to add the concept of an ecological self (Naess 1986, p. 82). What "resembles a definition of the ecological self" for Naess is simply: "The ecological self of a person is that with which this person identifies" (Naess 1986, p. 83). He thus agrees with relational views of the self, which emphasise the relation as the basis of a self but disagrees with the view that these relations have to be with other humans or with society at large. An ecological self is relational and encompasses relations with non-human animals, plants (Naess 1986, p. 82) and with inanimate nature. Because of the Naess' intention of including inanimate natural entities into the sphere of moral concern, he proposes to use the term "ecosphere" rather than "biosphere" because of the latter terms limitation with what concerns the science of biology (Naess 2005, p. 112).

Naess' account of the genesis of the self is comparable to Buber in particular and to the prevalent view in the psychology of personality in the last century. The primary state of the human being is "a vague network of relations" (Naess 1986, p. 88), which is only gradually substituted by the subject-object duality or as Naess puts it "the tripartition: subject, object, medium" (ibid.). Interestingly Naess associates the initial monism of the child with Bertrand Russell's neutral monism and takes a better understanding of the ecological self to be a move away from the dual or triadic conceptions of later life (ibid.). This diagnosis is not only relevant for a conception of self but directly for moral philosophy as it influences the perception of other creatures. Naess argues that the subject-object dichotomy is not a given but partly a consequence of training in that particular framework: "Education in industrial countries is strongly centred on a subject-object cleavage: Some traits of animals are real and objective attributes; others are said to be projected onto the animals. The latter traits are merely subjective. In practical life, the distinction is a plus, but it downplays spontaneous experience with its richness, intensity and depth" (Naess 1995, p. 283).

Another challenge involved in the development of the self as a moral agent is to attain "comprehensive maturity" (Naess 1986, p. 81). Naess understands comprehensiveness as "being mature in all major relationships" (ibid.), i.e., as a wide identification with all living beings, not only with some. It is possible, he argues that an individual might be mature in one set of relationships, e.g., with other adults, with family members, with members of her tribe, class or nation while displaying weak identification with others, e.g., nonhumans, children, foreigners etc. Descartes, Schopenhauer and Heidegger are examples of failures in comprehensive maturity, precisely because they are not capable of approaching animals, family

members and political others in the same mature manner they display in other relationships (ibid.).

A further assumption of Naess is that it is possible to actualise more or less of one's unique "potentials", a process he calls "self-realization" (Naess 1986, p. 82). This process contributes to a good life; it increases the "joy we experience in living" and even "the meaning of life" (ibid.). Naess' stand on the realisation of the particular potentials that a being has is thus comparable to that of Buber, who also argues that every being had a unique set of inherent potentialities, the actualisation of which is good.

Which beings can be thought of as falling under the addressees of Naess' supreme norm of self-realisation? Naess is vague about the borders, when he unhesitatingly counts animals and plants among such beings but argues that expanding the circle to include "a wider range of things dilutes the very concept of realization and Self. There is a limit here, but it is not definite, and the options regarding how to trace it are many" (Naess 1985, p. 196). Nevertheless, Naess' expanding circle is wider than Peter Singer's, since his understanding of the concept of interest is nearer to Schopenhauer's will than to that of (preference) utilitarianism: "The interest, in a broad, easily understandable sense, of a tiny plant to live and blossom is obvious, and under suitable circumstances, we act to serve this interest" (Naess 1992a, p. 296).

A subscription to this normative postulation carries along with it the need to study and question what such inherent potentialities might be; of the species in general as well as of an individual specimen in particular (Naess 1986, p. 86): "Animals and plants have interests in the sense of ways of realizing inherent potentialities, which we can study only by interacting with these beings" (Naess

1986, p. 86). The close bond between self and other leads to self-realisation being only possible as a joint venture. Subscribing to the subject-object dichotomy and considering the self to be an independent entity leads not only to a lesser self-realisation of the other but to that of the self as well: “Our self-realization is hindered if the self-realization of others, with whom we identify, is hindered” (Naess 1986, p. 82).

What exactly does Naess understand by “identification”? It is not the case that a person, who identifies herself with another entity, comes to believe that she is that entity but rather that “we ’see ourselves in others” (ibid.). Unlike Buber, Naess considers the concept of empathy to be central to identification and thereby to both the genesis of the self and its structure. A “paradigm situation involving identification [...] would be a situation that elicits intense empathy”⁶⁷ (Naess 1986, p. 83). Again unlike Buber, identification with the other is a prerequisite for compassion and solidarity (Naess 1986, p. 84).

The Link between Self and Morality

When considering the nature of the good life, i.e., the good, the increase of which can be considered as desirable, Naess positions himself against regarding the place of that good as the subjective consciousness of an agent. In a manner very similar to Buber, he considers the good – e.g., joy – to be neither in the subject nor in the object

⁶⁷ As with Buber’s childhood story with the horse, Naess’ example deserves to be quoted in full: “My standard example involves a nonhuman being I met in the 1940s. I was looking through an old-fashioned microscope at the dramatic meeting of two drops of different chemicals. At that moment, a flea jumped from a lemming that was strolling along the table. The insect landed in the middle of the acid chemicals. To save it was impossible. It took minutes for the flea to die. The tiny being’s movements were dreadfully expressive. Naturally, I felt a painful sense of compassion and empathy. But the empathy was *not* basic. Rather, it was a process of identification: I saw myself in the flea. If I had been *alienated* from the flea, not seeing intuitively anything even resembling myself, the death struggle would have left me feeling indifferent.” (Naess 1986, p. 83).

but as “a feature of the indivisible, concrete unit of subject, object, and medium” (Naess 1986, p. 94). Disregarding this indivisibility and postulating independent existence to the three elements of such an actual, concrete unit leads to a distortion of all three.

Naess calls himself a “student and admirer since 1930 of Gandhi’s nonviolent, direct actions in bloody conflicts” (Naess 1986, p. 89) and accordingly positions himself near to Gandhi’s ontology of the self. The ultimate goal or the highest good for Gandhi is self-realisation, where the self in question is not the ego of western philosophical traditions but Atman. Atman is a wide self, in which the opposition of ego and alter that corresponds to the opposition of subject and object is transcended (Naess 1986, p. 92). The link between Ahimsa (non-violence) as a principle of moral action and this conception of self lies in the connection of all beings through the relational structure of the self. In other words “if your self in the wide sense embraces another being, you need no moral exhortation to show care” (Naess 1986, p. 91).

Identification in a broader sense with animate and inanimate components of the environment can lead to a wider self, which acts naturally in a careful and concerned manner, taking the interests of other beings into account as part of their own self:

We need environmental ethics, but when people feel that they unselfishly give up, or even sacrifice, their self-interests to show love for nature, this is probably, in the long run, a treacherous basis for conservation. Through identification, they may come to see that their own interests are served by conservation, through genuine self-love, the love of a widened and deepened self (Naess 1986, p. 85).

This is not the only basis for Naess’ conception of morality however. Unlike Buber, he lays weight on the internalisation of ethical norms and habitualisation of moral behaviour. Based on Kant’s distinction between moral and beautiful actions (Naess

1993), he argues that it is essential to cultivate and promote inclination in “every aspect of socialization and acculturation and therefore also in the global ecological crisis” (Naess 1993, 138).

This link between the conception of self and a particular set of moral principles is valid only if it is true that a moral agent naturally cares for his self. Additionally, it needs to be shown why such an expanded sense of ethical egoism is acceptable.

CHAPTER 6

CONCLUSION

In this dissertation I have attempted to work in the cross-section of two areas of philosophical enquiry: The nature of the self and the conditions for effective moral agency. I have argued that these areas are not unconnected. Rather, how we conceive of the self has implications for how we conceive of moral agency and vice versa. An investigation into both these areas and their interconnectedness is also informed by the unique character of the self as an object of study, which can be accessed both from a first person view as well as a third person view. Martin Buber's moral ontology suggests that these two perspectives can be complemented by the second-person view, which considers the relation between a human being and the other as central.

My overall goal was not predominantly action-guiding as to argue for a particular normative moral stance. My aim was rather to critically investigate what must be true of an entity to be considered as a moral agent and whether a particular conception of self, i.e., the conception of self as relational is better equipped than some of its rivals for understanding and developing moral agency. With this goal in mind, I have drawn up a general account of the debate around morality and moral agency on the one hand and around self on the other. I have then gone on to discuss in detail two examples of philosophers who endorse a relational conception of self. In this concluding chapter I shall return to the leading question of the dissertation and consider what a conception of relational self adds to moral agency.

The first question is what sort of a conception of self had best be adopted. I have argued that although there are conceptual difficulties surrounding the term, its use is justified and useful. Nevertheless, a broad overview of some of the different conceptions of self in the history of philosophy was enough to demonstrate that views on both the nature of the self and what a particular conception of moral agency calls for have been and continue to be immensely diverse. My aim has not been to provide a comprehensive account of traditions, concepts and postulations around the theme of self but rather to concentrate on the aspects, which are relevant for moral agency. The philosophical landscape from this vantage point can be described along some few lines of thought. One line is ontological and concerns the question, whether the self had better be understood as an independently existing thing or rather as a conceptual tool to denote the totality of mental and/or behavioural states. The debate here is fraught with difficulties, since what one party may accuse the other to have postulated may be a straw man, as in the frequent portrayal of a wispy Cartesian ego.

I have exemplified the critique of such a substantial account of the self by referring to the positions of Hume, Nietzsche and Parfit. All three reject the idea that there is a subject behind the thinking and the doing. A second dimension of the discussion was exemplified by Sorabji's position that the self exists but not in the form of the *res cogitans* of Cartesian dualism. According to this view the self is the embodied owner of experience and necessarily so as the provider of diachronic connectedness as a condition for agency. The second line of thought thus affirms the existence of the self and discusses whether it is embodied or not.

A third line of thought likewise rejects substantial accounts of the self and develops alternative accounts, which consider it not only as being embodied but also

as necessarily embedded in socially and culturally contingent bases of meaning. I have considered narrative accounts of the self as part of this attempt, exemplified by Dennett's use of the term as a theoretical tool and further considered Bourdieu as an example of a sociological approach to the question of self. What is particularly noteworthy about such positions is that they localise the question of self in a network of collective discourses on the structure or even purpose of world, life, agency and axiology. This view has methodological implications, for if it is true that "In the domain of morality, as a lived phenomenon and as an area of inquiry, neither philosophy nor psychology nor social and political theory serves as the foundation for any other" (Flanagan 2009, p. 52) and if it is likewise true that the nature of the self cannot be investigated without taking the neurological, psychological, social and historical background into account, then a thorough investigation must involve a multi-dimensional, methodologically pluralist and inter-disciplinary approach. In the context of this dissertation I have only been able to draw a general framework for such an approach, the further elaboration of which remains to be pursued.

The question of which particular aspects of self are relevant for moral agency are also subject of scrutiny in psychology in the debate around the determinants of judgement and behaviour. In the cross-section of philosophy and psychology these are discussed among proponents of character theorists and virtue ethicist versus situationists. I have argued with Owen Flanagan that both aspects are to be taken into account and I have further argued that the Social Intuitionist Model (SIM) proposed by Jonathan Haidt is a useful and comprehensive model to work with. This is because the self is influenced by a great variety of factors; social as well as political, neurological as well as self-constitutional and it can only be adequately studied by taking this variety into account.

In this network of factors, it is expedient to distinguish a ‘sense of self’ from a ‘conception of self’. A sense of self is the phenomenological, first person view of one’s own self. I have argued that such a sense of self is closely related to the conception of self. The phenomenological view of what one’s self is, will influence one’s theory of self and vice versa. Assuming that many if not most people have a sense of their self as singular or unified, the prevailing conception of self as such can partly be explained by it.

Following conceptual clarifications around the conception of self and phenomenological approaches to a sense of self I briefly turned to the notion of the embodied self. If it is true that the self is not a Cartesian *res cogitans*, which exists independently of the body, then an investigation into the self must also involve its physical bases. In the case of the human self, this necessitates a biological and neurological approach, as well as a philosophical one. I have argued that studying split-brain patients or pathologies of the self is a good method to understand how a sense of self is created in non-pathological cases. I took the pathologies discussed to indicate that individuals are not neutral towards possible senses of self but tend to favour one, which is single or unified, both synchronically and diachronically. I argued that the frequency with which phenomena like confabulation and depression occur in cases of pathologies of the self can be interpreted as indicators for one of my main sets of assumptions: Firstly the idea that a normal development produces a sense of a single or unified self, secondly the idea that such a ‘normal development’ can only take place under certain neurological, psychological and social conditions and lastly the idea that aberrations are generally – but by no means necessarily – experienced as unpleasant.

This last factor has important implications for moral philosophy. If it is true, this means that it will cause a set of perceptions, judgements and behaviours bent on avoiding a sense of dissociation, however slight. Further indications for the existence of such a motivational aspect for the creation of a singular or unified sense of self besides confabulation and depression are distorted and self-enhancing memories, post-hoc rationalisations and self-deception. I have mentioned Buddhism and Schopenhauer's quietistic philosophy as examples for the idea that recognising the self as an illusion should lead to emancipation from its shackles, to the lifting of the veil of Maya. There can also be other conclusions drawn from the idea that the self is, if not quite an illusion then not quite what it seems. Much modernist thought deals with the idea of multiplicity and incongruity, not as a blemish of the self but as a celebrated element of the particular postulated *ars Vivendi*.

After endeavouring to gain some conceptual clarity about the usage of the central terms and the general framework of the problem, I have gone on to discuss Martin Buber and Arne Naess as two examples of philosophers who take a relational view of the self as a basis to argue for a particular moral stance. The link between Buber's philosophical anthropology and his ethics is to be found in his claims that human beings act rightly when they act from a unified *Wesen* and that the gradual unification of the various and conflicting parts of the soul happens only in encounters with others: „For the innermost growing of the self does not happen, as is often assumed these days, from the relation of the person to himself, but from that between one and the other, among people particularly from the reciprocity of realisation – from the realisation of the other self and the knowledge of being realised by the other in one's own self – in one with the reciprocity of acceptance, the affirmation and recognition and acknowledgement” (Buber 1950, p. 423, translation B.S.).

The link between Naess' conception of self and his ethics can in a certain sense be interpreted as egocentric. Naess' position on the nature of the self is comparable to that of Buber in that it is relational. This is where the similarity ends however. For Naess "The self can be said to comprise what one identifies with" (Naess 1982, p. 248). For Buber, the alterity of the other remains unalterable and central. The other, although affirmed and acknowledged in an I-You relationship, is not appropriated by being included in a wider sense of the self. She does not become part of the *oikos*; on the contrary, remaining alien is a decisive element of encounter. Naess' understanding of the genesis of self as wider identification with animate and inanimate nature on the other hand, involves a sense of belonging between the I and the other. Naess does not simply state this as an element of the self. Such identification is normative and Naess considers it to be a sign of maturity: "The person, I suggest, who is 'all-round' mature cannot avoid identifying with every living being – seeing himself or herself in every being" (Naess 1982, p. 248).

Naess' idea of an extended self as the product of a wider identification with the animate and inanimate environment is the foundation for an egotistical argument: "We naturally and spontaneously care for our place and seek to protect it. For this we do not need a moral axiology, a set of rules or enforcements held over us to force us to act" (Drengson 2008, p. 37). It is important to distinguish this process as a postulation of descriptive moral psychology from a normative prescription. The latter would collapse into a type of egotistical morality, which could be distinguished from more traditional egotistical normative theories only by the wider compass of the ego.

This is the weak spot of Naess' argument that the recognition of the narrow ego as an illusion will lead to the espousal of a wider self, which considers the distinction between the I and the other likewise as an illusion and incorporates the

latter as a natural part of the own self. If we subscribe to the principle of equality, there is no evident reason why one's interests should have precedence over the interests of others simply on account of being one's own. The widening of the ego as the sole or even main reason to practice moral concern for other entities involves the danger of collapsing into familiar types of chauvinism. I therefore argue that Naess' argument from self to morality remains problematic without a separate argument why an expanded self-interest should be a morally acceptable position to take.

Unfortunately Naess hasn't written much on the link between his conception of self and his moral philosophy but Buber elaborates on many aspects of moral agency and normative ethics. His figure of the human being is neither the atomistic individual drawn and criticised by communitarianism, nor the figure of the faceless member of a community drawn and criticised by individualism. His philosophical anthropology situates the person in a wide and differentiated net of factors that contribute to the self of the moment. The question of moral agency is therefore linked to this net of factors, which, on the whole, support a moral philosophy based on particularism and situationism.

I take the comparison of Buber and Naess to demonstrate two things: Firstly, although they both argue for a similar ontology of the self as necessarily relational, they draw different consequences from this starting point. For Buber the other remains alien and the moral situation is done justice to if this fact is acknowledged. For Naess, the idea that the other is alien is an illusion and a mature self is wide enough to encompass it. Their divergence from a comparable common starting point gives us reason to argue that the relationality of the self does not necessarily force a certain normative stance on us. Buber and Naess give us good reasons to endorse the

idea of the self as necessarily relational, while enabling us to draw our own conclusions for moral philosophy.

A comparison of Buber and Naess further provides us with different ways to consider the position of non-human moral patients. There might be historical reasons behind the fact that while Buber acknowledges nature in general and non-human living beings in particular as possible others in an encounter, he does not himself elaborate on the moral implications involved. Naess on the other hand, involved in the environmental movement from its very beginnings is very sensitive towards the plight of non-human natural entities at the hands of human moral agents. While neither Buber nor Naess argue for an expansion of the circle of moral agents to encompass non-human animals, I take their positions about the role of such entities in the creation of a moral situation as successful in establishing the need for accepting responsibility towards moral patients, whether as the alien other or as part of one's wider self.

I take these two points to support the view that both Buber and Naess contribute to the development of a procedural approach to moral philosophy, which remains largely neutral towards the set of adopted values and allows the development of a variety of singular approaches. In Naess' case, these can take the form of individual or collective ecosophies, which may or may not converge on particular rules, decisions or actions (Naess 2005).

This endorsement of a plurality of methods and concrete decisions does not imply that values can be chosen arbitrarily. Buber for example is explicit about his view that moral values are not invented but discovered (Buber 1953a, 554). I partly follow Buber in assuming a stance of moral realism but contrary to him I am more wary of absolutes. I maintain a less ambitious position and take moral judgements or

arguments to be more or less successful in being impartial, consistent, complete or critical while suspending judgement on the deeper ontological question concerning moral facts. This is the one of the characteristics of what I call a procedural moral theory, the point of focus of which is the determination of which factors make up a moral situation and what conditions need to be met to make moral judgement and moral decision better in the sense above.

Buber, as we have seen, argues from the nature of the self as relational to a moral theory, which is developed around the *Wesen* of human beings. I interpret this aspect of Buber's thought as a problem of authenticity. Morally right action involves authenticity to one's own *Wesen* and acknowledgement of that of others. The main elements, which are involved in the actualisation of the situation as constituted by the authenticity and relationality of those partaking of it are dialogue, inclusion (*Umfassung*), the capability of perceiving the demands of the situation and the responsibility one undertakes towards it. The situation in question is anchored in daily life as the realm of morality and it can only be realised if the participants are authentic.

Advocating authenticity and conceiving of authenticity as a central concept of moral philosophy opens up some questions in connection with the relational view of self. If the self is indeed relational and it makes only sense to speak of it as part of a relation, whether as part of the pair I-It or I-You as Buber postulated or as a wide self in relation with human and non-human entities as Naess conceives it, then authenticity seems to be a paradoxical postulate. If it is true that there is not as much of a 'true' self, not as much synchronic or diachronic identity as our sense of self leads us to believe and if it is rather the situation along with its 'others', which predominantly determines the momentary self then what is there to be authentic to?

The term authenticity implies that there is an entity, which can be more or less near to what its 'true' *Wesen* is.

There are two possible answers to this question. One includes a rejection of Buber's conviction that there is a divine way of the world, of whose actualisation every human being is responsible by being authentic. Belief in a master-plan of the world, its inhabitants and their relations involves belief in a planner, which is an ontologically extravagant position to take. A secular version of this approach substitutes divinity with nature and argues that to be authentic is to be natural. Despite avoiding the postulation of a divine planner, the concept of nature itself is problematical in that it involves a questionable abstraction, especially when nature is marked off against culture. Furthermore, even assuming that either one or both of such postulations of a primal and true state of the world is accurate, the move to normativity must yet be argued for.

The second answer makes less substantial assertions about metaphysics. Authenticity in the framework of relationality can also be interpreted as the capability and the willingness to be involved in the situation as an element leading to its actualisation rather than as an independent subject who deals with it. Naess' account of the human tendency of considering the self to be separate and independent as well as both Naess' and Buber's critique of the traditional subject-object dichotomy can therefore be interpreted as part of the failure to be authentic. Authenticity should be understood not as the correspondence to a set of essential qualities but rather to the state of contributing to the realisation of a relation, which is as little impeded by the sense and conception of the self as an independent and persisting being as possible. Authenticity understood like this becomes central for

procedural morality, not least because it avoids patronising or instrumentalising the other.

Authenticity requires that the moral agent must be attentive to the demands of the situation; she must not allow herself to be led by principles or calculations but rather allow herself to be moulded by the other and the situation. She must practice being in I-You relationships instead of I-It relationships. Such openness and affirmation does not imply that the I vanishes in the other. It is also not the case that the I-You has precedence over the I-It and that the I-It should be shunned as much as possible. Buber is realistic about the nature and capabilities of ordinary human moral agents: “It is not possible to live in the bare present. Life would be quite consumed if precautions were not taken to subdue the present speedily and thoroughly. But it is possible to live in the bare past, indeed only in it may a life be organised. We only need to fill each moment with experiencing and using, and it ceases to burn. And in all seriousness of truth, hear this:⁶⁸ without it man cannot live. But he who lives with it alone is not a man” (Buber 1958, p. 34).

The dialogical is in close relation with the I-You relation but there is no cognitive access to it to enable it to be taught or practiced as for example mathematics could be. In this, Buber’s dialogical ideal is related to the virtue ethicist example of the bright young teenager, who will necessarily fall short of responding in a moral situation like some older person who has acquired more *phronesis*. The only way of retaining, cultivating or developing the capacity for dialogue is the experience of dialogue itself (Krone 1993, p. 142). This is relevant for moral education and emphasises the importance of the life of the educator as a model and of the other in genuine I-You moments. Nevertheless, we must remain careful and

⁶⁸ What Smith translates as “hear this” and Kaufmann translates as “listen” (Buber 1970, p. 85) is actually “Du” in the original. Buber addresses the reader directly in this passage with a *You*.

take into account the phenomenon of age adversely affecting the capability to answer to the demand of the moment. Age makes us neither rigid nor wise necessarily nor is youth a guarantee for more attentiveness towards the situation and the other.

Procedural ethics must work with psychologists and educational scientists to develop a realistic account of how such skills are acquired or lost.

Since procedural ethics dwells on the skills which make it more likely for a moral agent to respond to the moral situation in a more focused and authentic way, it can gain from the insistence that the realm of the moral is everyday situations. For our purposes, Buber's call for realist particularism can therefore be translated into a critique of hypothetical thought-experiments in ethics in the style of Trolley-problems. Buber lived before the time that discussing such problems became common. Nevertheless, his understanding of what morality is and what it should be expected to do runs contrary to the assumption that such thought-experiments are able to help us in anyway. Buber might have even regarded them as dangerous, since they imply that it is possible to abstract in such a way from the complexities and singularities of the genuine situation, that many of the morally relevant aspects of a case are disregarded in favour of abstract calculation or the application of principles. A passage of Buber's about the need for responsibility and awareness in every situation backs the view that his attitude towards such thought experiments would most probably have been negative:

Each concrete hour allotted to the person, with its content drawn from the world and from destiny, is speech for the man who is attentive. Attentive, for no more than that is needed in order to make a beginning with the reading of the signs that are given to you. [...] For the attentive man would no longer, as his custom is, 'master' the situation the very moment after it stepped up to him: it would be laid upon him to go up to and into it. Moreover, nothing that he believed he possessed as always available would help him, no knowledge and

no technique, no system and no programme; for now he would have to do with what cannot be classified, with concretion itself (Buber 2002, 19).⁶⁹

I agree with Buber that the importance of attentiveness towards the situation cannot be overestimated. I nevertheless believe that thought experiments in ethics have their place and justification. Practicing authenticity and attentiveness and practicing intellectual speculation can be complementary moral pursuits rather than conflicting ones. There is no question that the attempt to isolate some specific factors in a given situation by making up hypothetical cases, ignores many of the factors that would be given in a real situation. This is both method and aim; for the purpose of such thought experiments is to focus on one specific factor or the other precisely in order to evaluate its position and importance in moral deliberation. Practicing to concentrate on one particular aspect of a morally relevant situation can moreover facilitate it for the moral agent to perceive more aspects of the moral situation to take into account. Like a conductor who learns to distinguish different instruments of his orchestra partly by abstracting them from the total musical constellation and studying them separately, hypothetical moral thought experiments can contribute to the skill of effective and comprehensive moral agency. I suggest that the decisive point is not whether some aspects of the situation can be abstracted and considered separately but rather the unfortunate tendency of coming to consider one such element as either the sole or the most relevant factor.

Buber's emphasis on the uniqueness of the situation is also mirrored in his demand for pluralism. His brand of pluralism is a peculiar one however, since it is

⁶⁹ This is related to the specifically Buberian brand of *Kulturkritik*, for he claims that "the whole apparatus of our civilization is necessary to preserve men from this attentiveness and its consequences" (Buber 2002, p. 19). Similarly, the person who has been attentive can still choose to remain in his habituated patterns of response or she can choose to remain silent, which Buber also diagnoses as an attitude that is "characteristic of a significant type of the age" (Buber 2002, p. 19). Buber's critique of the morally failing unresponsive person is thereby related to his *Kulturkritik* in general.

directly connected to his understanding of the world as creation. Every individual of any species has a unique position in this creation and has a contribution to make that can be made only by her.⁷⁰ Correspondingly, every individual has to follow her own path; meaning that there is no single, universally valid way of the good life. This call for plurality and individuality is twofold. It is a sign of a successful life that an individual maintains the attentiveness and open-mindedness to make informed decisions about her life choices, independently of reigning life schemes of the collective that the individual belongs to. It is at the same time part of a broader attempt of humanity to embrace diversity (Krone 1993, p. 129).

I have already dwelt in more detail on the evaluation of Buber's moral ontology. The broader question I am concerned with now is what his brand of moral philosophy has to offer and whether I am right in taking it as a type of procedural ethics. Buber is critical of the role that reason has been given in moral philosophy and pleads for a more comprehensive account of the moral situation. If we take it to be true that emotions and intuitions as well as heuristics, habits and automatic responses play an important role in moral judgement and moral behaviour, then Buber's insistence that attentiveness is one of the key elements of successful moral agency can be reformulated as the necessity for a moral agent to decide, whether the situation she is faced with requires such attentiveness. In other words, the first decisive moment in moral agency is the task to evaluate, whether the situation has morally relevant aspects or not. If the situation is considered to be a moral one, then following Buber, the moral agent must endeavour to be as attentive as possible to all relevant factors.

⁷⁰ „Jeder hat eine in Raum und Zeit ausgesparte Sphäre des Seins, die ihm zuteil ist, durch ihn erlöst zu werden“ (quoted in Krone 1993, p. 67)

If we choose to adopt this approach, we are required to distinguish between two moments of alertness, especially if we subscribe to Putnam's criticism that such attentiveness is too demanding to be realistically expected of ordinary moral agents. The first moment of what I shall call procedural attentiveness is necessary to decide whether the second, morally substantial moment of attentiveness is needed. The objection that such attentiveness is too demanding to be humanly possible could apply to both. In fact, to distinguish between procedural attentiveness and substantial attentiveness enables to answer this objection in part. The alertness required to distinguish between moral and non-moral situations is arguably less demanding than the attentiveness needed to decide how to respond to a moral challenge. It is possible to argue that the main hindrance to such basic alertness is not so much the mental energy required as rather a lack of familiarity with moral thinking, conceptual clarity and habit. Another way to respond to Putnam's criticism is the adaptation of Mill's answer to the lack of time argument against utilitarianism. This would take the opposite position and argue that an ordinary moral agent has enough familiarity with moral thinking, conceptual clarity and habit to respond appropriately to most moral situations.

Whether or not Buber would subscribe to such an adaptation of his concept of attentiveness, I argue that it is central for a procedural approach to morality. A further point is concerned with the question of the type of situation one finds oneself in. Being attentive to the demands of the situation is a must in order to act rightly but one does not often find oneself unawares in a given situation. On the contrary, a moral agent has a certain amount of power over which situation she chooses to be in. General decisions about where to live or what profession to choose will make it more or less probable that some situations occur rather than others. A moral agent who is

faced with a sick stray cat will have to take that into account according to Buber but how are we to evaluate someone who chooses to work in a sanctuary as opposed to someone who, perhaps consciously, avoids a lifestyle in which she will be faced with such morally challenging situations? Buber quotes the scriptures as God asking from the faithful to “be whole!” (Buber 1963, p. 621) and the wholeness in question would include how the cat will fare at the agents hands. Another slogan that I think could be ascribed to him would be “be there”. It seems that adapting Buber’s focus on situations to a procedural moral theory must be complemented with more work on more general choices of life-style.

I suggest further that it is also worthwhile to follow Buber in according dialogue a more central position in moral philosophy. There remains the open question however, whether dialogue ensures recognition of the other in a more or less accurate way. If we must accept the dictum that human beings are prone to self-serving biases, self-deception, flattering distortions of memory and similar, disempowering tendencies, then dialogue may not only be a way of affirming the other but might involve the danger of individual and collective deception. Such a danger is exemplified in the tendency among friends or in communities of establishing patterns of perceptions, judgement and action, which are mutually affirmed. An investigation into the procedural elements of morality therefore also involves a close scrutiny of such tendencies and possible techniques of forestalling them.

Buber prompts us to question aspects of moral philosophy that are traditionally in the forefront of discussions and proposes new alternatives. We have to consider, whether these alternatives are legitimate subjects of philosophical enquiry. Is it possible that there are aspects about morality (or more generally about

reality) that cannot be grasped by moral philosophy? Buber argues that there are and that philosophy in general or ethics in particular can take us only so far. If he is right, how do we approach such aspects? If the realm of the I-You, if the alterity of the other as my You is both as central as Buber argues and as impossible to conceptualise and be grasped by philosophical methods, how can we try to relate to them? This question does not imply that there might be metaphysically suspicious areas involved in morality, more akin to spirituality than to philosophy. It concerns metaphysically more respectable areas, which nevertheless are not as open to rational scrutiny as traditional moral concepts aim at being. How, for example, would we have to approach moral philosophy, if it had a musical element? Philosophy and psychology can only investigate music to a certain degree. A residue is condemned to remain inaccessible by philosophical and/or scientific methods and following Buber, an analogous case can be made for morality.

Buber's methodological approach or rather his non-approach, his resistance to systems is in congruence with his radical situationism and the constellation of the self as the result of the interaction between the I and either It or You. I suggest that it is nevertheless not necessary to reject systems altogether. The difference between Naess and Buber in this regard is exemplary of this point. Naess praises Spinoza for his systematic ethics and considers "the extreme consistency and tenacity with which consequences, even the most paradoxical, are drawn from intuitively reasonable principles" as a "major virtue" (Naess 1973, p. 131). While for Buber, the rejection of both systematic grand theories and the idea of an 'irrelational' I are based on his conviction that the uniqueness of a situation does not allow of reductive abstractions, Naess considers the very rationality, coherency and consistency of a set of inferences as a virtue. Independently on what the right position towards (moral) systems to take

would be, this methodological difference between the two authors demonstrates that the conception of the self as relational and interdependent within a wide set of environmental entities and events is not necessarily based on a certain methodological position. For the development of a procedural moral philosophy this provides the possibility of a variety of methodologies, which may include more systematical approaches than Buber thinks is permissible.

As a further element of such a situational, particularist, relational and procedural moral philosophy, Buber's anthropocentrism can also be taken as a challenge to develop a method of differentiating moral agents from moral patients and of determining their respective positions in a moral situation. Although Buber focuses on philosophical anthropology, the picture he draws of the moral situation allows an interpretation of non-human beings as morally relevant, albeit in a hierarchical manner borrowed from Chasidic tradition: "I find, by the way, that our relationship to the domestic animals with whom we live, and even that to the plants in our gardens, is properly included as the lowest floor of the ethical building. The Hasidim even see it as beginning with the implements of work." (Buber 1964, p. 28). Naess is also avowedly willing to consider non-human beings as morally relevant. I argued however that the position of the moral patient is precarious or instrumental in both espousals of a relational self as the focus of moral enquiry. Buber remains open to the criticism of anthropocentrism while Naess' plea for taking the wellbeing of animate and inanimate into account has elements of egoisticism. Neither Buber nor Naess provide a satisfactory account of moral patiency in the context of relational and procedural ethics.

Nevertheless and on the whole Buber's moral ontology and Naess' ecosophy provide us with useful tools to develop a procedural moral philosophy, which

remains ontologically cautious, which favours a pluralist approach, which focuses on a variety of elements of the moral situation simultaneously and which develops techniques to increase the chances of being a more successful moral agent in a given situation. Such a conception of moral agency also requires a model of the self to accommodate it. I have argued that among the variety of conceptions of self, those who consider the self not in a substantial way but as dynamic constellations of centres of mental activity, with their individual determinations seem to be more in congruence with the current state of psychology and neurology. I suggested to call these centres ‘monads’, because they can be considered as the smallest independent unit of mental life, without being otherwise similar to monads as Leibniz envisioned them.

The question may suggest itself whether monads are not identical to modules, which are increasingly losing their minority-status in debates about human beings as cognitive systems. Dan Sperber considers modules to be functional sub-systems, which are subject to evolution, which can modify themselves according to experience, which have domains proper to them, about which they inform the organism, which have their own procedures and their own database and which can affect each other only indirectly. Furthermore, modules are “neurologically distinct devices”, which may but must not necessarily correspond to singular brain locations (Sperber 2005). If we follow Sperber in thinking that mind-brains are best viewed as articulations of specialized modules” (Sperber 2005, p. 53) don’t modules appear to be identical to monads?

I have used the concept of a monad as a theoretical tool to dwell on not so much the mind-brain but rather the conscious and unconscious mental life of an agent. Following Buber I use the term ‘mental’ not with the intention of ignoring the

body of the agent but to emphasise that what is in the focus of my interest is not purely the third person view of human behaviour but also the first person view, the sense of oneself as a moral agent and the second person view, which accords the other a central place in the self. Monads share a number of characteristics with modules but are not identical to them. Like modules, monads are sub-systems but unlike modules any such a sub-system can theoretically function as the sole unit of consciousness at a given time. Like modules, monads are embodied, but unlike modules, I do not claim that they correspond to or are identical to particular neurological modules. Monads are psychological constructs rather than tangible neurological categories, even if they might in time be demonstrated to correspond to neurological units, analogous to what Fine calls “schema”:

Cognitive psychologists think that just about everything we learn about the world is neatly tidied away into a schema. I like to think of a schema as a big bed full of slumbering brain cells. All the brain cells in the bed represent a different part of the schema. So, for example, in the schema for dogs you'll find brain cells that – when active and awake – point out that dogs have four legs. Then there are neurons that hold the information that dogs bark, neurons that remind you that dogs have hair, and all the neurons for just about everything else you know about the concept of dogs. And they're all tucked up in the same bed (Fine 2005, p. 131).

Thus, we return to one of the main questions of the thesis, i.e., the question of how a sense of a unified self is created and what implication this has on moral agency.

Sorabji rejects the idea that there is a faculty which is responsible for bringing the innumerable perceptions and volitions together: “If there is unity in one’s self-awareness, the unity is supplied by the single owner of that awareness, not by the owner’s using a single faculty” (Sorabji 2006, p. 260). He believes that there is a unitary and single self, rather than an ever-changing stream of consciousness without an owner and he argues from evolution that “survival requires us to be aware of it, so that it is not surprising that natural selection has structured us to be so aware”

(Sorabji 2006, p. 261). I return to this point, because it is exemplary for the position that I have argued against:

- Self is, by its nature, unified (Third person view).
- Conception of self is conception of self as a unified entity (First person view).
- Conception of self is conception of self as a unified entity *because* self is, by its nature, unified.
- It is necessary for effective agency both that the self be unified and that there be self-awareness of this.

I have argued that these assumptions are problematical. Regarding self-awareness, phenomenological reports of both normal senses of self and a variety of pathologies of the self demonstrate that a sense of the self as unified is by no means the only case, even if it might be the most common. Especially pathologies indicate how the brain deals with an immense number of bottom-up and top-down processes of cognition, perception or volition and organises these to create a sense of singular stream of consciousness. Let us assume that it is true that the brain fabricates the self and also that a number of social structures contribute to conceiving of the self as an entity with a past, present and future. As an explanation of how a conception of self as unified comes about, these processes would explain precisely that and would not suffice to show that the self is therefore not ‘real’ but an ‘illusion’. I reject the first three propositions of the traditional view paraphrased above. Nevertheless, I do not believe that such a rejection is tantamount to a proof that the self is an illusion and a construct. There might be other ways to demonstrate that it is – if, indeed, it is – but

assuming that a rejection of these three propositions is sufficient, amounts to a genetic fallacy.

But the real question concerns Sorabji's postulation that effective agency requires both a unified self and an awareness of that self as such. If this were true, moral agency would have to be denied to individuals who do not have or are not a self – depending on whether we want to think of one as identical with or as the owner of one's self. It would also have to be denied to individuals who do not conceive of themselves as unitary – regardless of whether their self is in fact unified or not. This seems far too narrow a conception of who can be considered as moral agents. It requires not only individuals to have a first-person sense of their selves as singular or unified, it also requires them to have this sense for the right reason, i.e., for the alleged fact that they do have singular or unified selves. Apart from the fact that there are no conclusive reasons to subscribe to Sorabji's argument, such criteria reduce the possible candidates for moral agents to a minimal part of human beings, if any.

Another sub-question of the implication of how the self is structured (3rd person view) and of how individuals conceive of their selves (1st person view) on ethics regards not moral agency but the position of moral patients. Here the question of singularity, unity and ownership is not concerned with how (moral) action can be modelled but rather with what elements of moral patiency are related to self and what this relation requires. Sorabji suggests that without ownership there are no sufferers and no beneficiaries (Sorabji 2006, p. 273); i.e., without the individual understood as the owner of her self, there would be no moral patiency. It is not clear why suffering or being benefited should require a conception of self as ownership. But if it is the case that moral patiency is best understood via suffering and/or interests, then

diachronic identity does seem relevant to moral patiency as including past and future selves. One reason for this is the concern of my present self for my future self or the applicability of moral evaluations like remorse or regret for the attitudes or actions of my past selves. Another reason is the fact that, similar to many types of actions, many types of events of which the moral patient is the recipient take time and involve many constellations of the self over that amount of time. This is related to the Aristotelian point that it is impossible to call a life as passed well before the individual dies and even after their death some events – like the lives their children lead or the course their projects take – could arguably contribute to the success or failure of that life.

To recapitulate; I have separately argued for three elements of self and morality: Firstly, I have argued for a monadological model of the self. Secondly, I have argued that relationality is an essential quality of the self. And thirdly I have argued that a procedural approach at morality is needed, even if by itself it may not be sufficiently comprehensive to claim to replace more traditional moral theories. If I am right in adopting these three elements, then the picture of the moral self that emerges is characterised by the following: The moral self is an aspect of the self understood as an event. An indefinite number of monads of varying precedence make up a unique constellation of this self at a given moment. These monads are necessarily relational as is the overall self as a product of their momentary constellation. Whether the self of a given moment includes the element of a barking, hairy dog as in Fine's example of a schema, depends on whether that monad is activated, whether it is conscious and in what relation it is to other monads.

Leibniz' monadology is not only a good model for the sort of self that I have in mind; his understanding of this world as the best of all possible worlds is also

helpful as a metaphor for the tension between the unity and the diversity of the self. It might make sense to argue that neither complete unity (regardless of whether most people welcome it or not) nor great diversity are conducive to moral agency. The "best of all possible selves" might in this sense be the one, which combines the highest degree of unity and variety compatible with each other.

Outlook

I have been arguing for the overall thesis that there are good arguments to work with a monadological model of the self and that this model enables us to further develop a moral theory, which is relational and procedural. Part of the work that would be involved in such a project concerns moral development and moral education. Much work has been done on the cognitive development of children and it has been argued, most famously by Jean Piaget and Lawrence Kohlberg that moral development runs parallel to the acquisition of cognitive faculties. In later years, the role of emotions and intuitions in moral judgement and moral behaviour has received increasing interest. The idea that "a developmentally ordered progression exists in moral development, in which conditioned behavior precedes explicit thought" (Baird 2008, p. 324) is in accordance with the stages of moral development postulated by Kohlberg, in which moral values are first internalised along parental sanctions in childhood. What has changed in the study of moral development is that today the autonomous, sovereign, post-conventional moral agent is no longer considered to be the typical figure of later stages of moral development. Baird, for example, argues that "the development that takes place in adolescence, namely, the integration of intense visceral emotion with social cognition, is essential for a fully developed

moral reasoning that functions pre-emptively, with minimal cognitive effort” (Baird 2008, p. 324).

A further development of relational and procedural ethics would have to take such a psychologically realistic account of moral development into account. Its anthropological premise that the human self is monadological requires interdisciplinary work on the development of the moral self in a more differentiated manner than the traditional models of moral development as being dependent on and nourished by ever-increasing cognitive skills.

But apart from these preliminary steps, there are further perspectives in a comprehensive enquiry into the legitimacy of the criteria of exclusion and inclusion from the moral community that need to be taken into account. It seems clear that a comprehensive discussion of both agency and patiency in procedural moral theory must involve the sociological and political aspects of the exercise of free will, of coercion, domination, power and collective action. Paternalistic policies for example, in which people are excluded from being active agents in the political process by open discouragement and insufficient training of active citizenship in the educational system leave too little space for moral agency to flourish.

Further open questions concern moral development and education. One of these questions concern the moral status of different people endowed with different grades of empathy, sympathy or care as affective states influencing moral action or a concern for equality, justice or right. How ought one to treat people displaying varying degrees of moral merit, if the meritocratic approach of a moral of reciprocity leads to injustice? And how must moral education be organised in order to

accommodate varying degrees of readiness, inclination and aptness in moral thinking?

To continue, if I am right in supposing that the traditional understanding of the moral self as an independent, rational decision-maker has been successfully challenged then the most promising candidate to replace it, i.e., the relational theory of self opens new questions about responsibility. If it is true that „[a] human being alone is an impossibility, not just de facto, but as it were de jure“ (Taylor 1985, p. 8) then it seems that accountability and responsibility also lose their accustomed focal point, the individual. If the self is relational, then there must be a more systemic relational point of responsibility rather than the agent by herself. It seems bewildering to remove the moral agent from her position of responsibility and yet, if the relational theory of self accommodates more facts about agency than the traditional one, then the question of responsibility must be a further line of enquiry to be pursued. Similarly open to challenge is the moral status of habits, habitualised actions and automatic actions.

The final open question to be pursued concerns the amoralist, an anti-hero difficult to fit into any moral theory. The amoralist is in the centre of a number of questions concerning moral development, moral character and moral motivation and poses a challenge to all three of the canonical normative theories. A moral agent is a figure we need in our enquiry into the moral community, both directly as an agent as well as indirectly as the subject treating moral patients. However, there seems to be no clean way of accommodating the amoralist; the person, who is endowed with the potentialities of a moral agent but who, of his own free will, refuses to regard anything as morally binding or who denies the possibility of moral philosophy to begin with. This character seems to be the greatest challenge to any moral theory for

there is no way known to me of approaching the amoralist from any starting point, which is not moral itself.

Another question that could be followed up concerns the discussion of the nature of selves. While taking a survey of the various conceptions of the self, it appeared that different as these were, most seemed to assume that there is one type of an entity, which all individuals partake of. Whether these individuals are only human beings or also include non-human animals, they are thought to be characterised by having the type of self that a particular view in question postulates. Not even the Buddhist view of no-self is an exception, as it also assumes that all individuals fall under this type.

There remains another view to be considered however, which could tentatively be called a pluralist notion of self: Is it possible that the ‘self’ is different from individual to individual? This is not the familiar question of personal identity – the question of what makes me me and what distinguishes you from me – but the question whether different types of self might in fact be distributed among a given population. One could then claim that every individual has a self like every individual has a body but that there are different types of selves as there are different types of bodies. While the self of an individual with a particularly good memory and a strong relationship with her past selves might be nearer to the type of a narrative self as Dennett or Plutarch postulate it, another individual with less strong memory and less interest in former or future selves like Strawson might have a self that is structurally more similar to the pearl-model that he postulates. Some characteristics of the self might be necessary – like the physical extendedness of the body, whatever type it might else belong to. I suggested that the relational nature of the self is an “essential” characteristic of the self but argue that the monistic quest for *the* nature of

the self might otherwise be misguided. The possibility that various conceptions of selves do not compete but are actually descriptions of different types of selves should be investigated further.⁷¹ If I am right in arguing that all these types of self are relational in varying degrees, then a relational, procedural morality such as I advocate must develop by taking this variety into account.

A comprehensive discussion of the moral self, moral agency and the moral community must consider all these aspects of morality critically, inter-disciplinarily and anti-parsimoniously. I would like to adapt Haidt's (2008) methodological wish of psychological theories to moral philosophical ones and demand of them to "have the *optimum* amount of complexity, not the *minimum* that a theorist can get away with" (p. 205).

⁷¹ A hint of this idea is already present in studies of unusual conceptions of self, such as those of autistic people as is indicated in the following description of two autistic children: "Their reduced sense of self removes from them pride, embarrassment, shame, humiliation, vanity, feeling sorry for what they did, ambition, emulation, or the idea of reward [...] [They] are not aware of a *past* and *future* self" (Sorabji 2006, 29).

REFERENCES

- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, D.C: American Psychiatric Association.
- Anderson, A. (2008, November). Life on Man. *The Economist*, 150. Retrieved from http://www.economist.com/node/12494708?story_id=12494708
- Aristotle (1996). *Nicomachean Ethics*. Hertfordshire: Wordsworth Editions.
- Baggini, J., & Fosl, P. S. (2007). *The Ethics Toolkit. A Compendium of Ethical Concepts and Methods*. Oxford: Blackwell Publishing.
- Baird, A. (2008). Adolescent Moral Reasoning: The Integration of Emotion and Cognition. In Sinnott-Armstrong, W. (Ed.), *Moral Psychology. Volume 3. The Neuroscience of Morality: Emotion, Brain Disorders, and Development*. Cambridge, Massachusetts: MIT Press.
- Barnes, B. (2000). *Understanding Agency. Social Theory and Responsible Action*. London: Sage Publications.
- Bartky, S. L. (2002). Agency: What's the Problem. In Bartky, S. L., "Sympathy and Solidarity" and other Essays. Maryland: Rowman & Littlefield Publishers.
- Bem, D. J. (1972). Self Perception Theory. In Berkowitz, L. (Ed.) *Advances in Experimental Social Psychology, Volume 6*. New York: Academic Press.
- Bentham, J. (1996). *Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press.
- Bourdieu, P. (1982). *Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft [Distinction: A Social Critique of the Judgement of Taste]*. Frankfurt am Main: Suhrkamp Verlag.
- Bourdieu, P. (1985). *Die biographische Illusion [The biographical Illusion]*. In: Bourdieu, P. *Praktische Vernunft. Zur Theorie des Handelns [Practical Reason: On the Theory of Action]*. Frankfurt am Main: Suhrkamp Verlag.
- Bourdieu, P. & Wacquant, L. (1996). *Reflexive Anthropologie [An Invitation to Reflexive Sociology]*. Frankfurt am Main: Suhrkamp Verlag.
- Buber, M. (1919). What is to be Done? In: Buber, M. (1957). *Pointing the Way. Collected Essays*. New York: Harper & Brothers.
- Buber, M. (1922). Religion als Gegenwart [Religion as Presence]. In: Reichert, T. (1996). *Buber für Atheisten. Ausgewählte Texte. [Buber for Atheists. Selected Texts]* Gerlingen: Lambert Schneider Verlag.
- Buber, M. (1926). Education (Rede über das Erzieherische), In Buber, M. (2002).

- Between Man and Man*. London and New York: Routledge Classics.
- Buber, M. (1930). Zwiesprache [Dialogue]. In *Das dialogische Prinzip*, Gütersloh: Lambert Schneider/Gütersloher Verlagshaus.
- Buber, M. (1930b). Dialogue. In Buber, M. (2002). *Between Man and Man*. London and New York: Routledge Classics.
- Buber, M. (1935). Bildung und Weltanschauung [Education and Worldview]. In Buber, M. (1964a), *Reden über Erziehung [Speeches on Education]*. Heidelberg: Lambert Schneider Verlag.
- Buber, M. (1938). What is Man, in Buber, M. (2002). *Between Man and Man*. London and New York: Routledge Classics.
- Buber, M. (1939). The Land and its Possessors. An Answer to Gandhi. In Herberg, Will (Ed.). *The Writings of Martin Buber*. New York: Meridian Books.
- Buber, M. (1939a). The Education of Character. In Buber, M. (2002) *Between Man and Man*. London and New York: Routledge Classics.
- Buber, M. (1947). Der Weg des Menschen nach der chassidischen Lehre [The Path of Man along the Chassidic Dictum]. In Reichert, T. (Ed). *Buber für Atheisten. Ausgewählte Texte. [Buber for Atheists. Selected Texts]*. Gerlingen: Lambert Schneider Verlag.
- Buber, M. (1948). Das Problem des Menschen [The Problem of Man]. In Buber, M. (1962) *Werke. Erster Band: Schriften zur Philosophie [Collected Works. First Volume: Writings in Philosophy]*. München and Heidelberg: Lambert Schneider Verlag.
- Buber, M. (1950). Urdistanz und Beziehung [Primal Distance and Relation]. In Buber, M. (1962) *Werke. Erster Band: Schriften zur Philosophie [Collected Works. First Volume: Writings in Philosophy]*. München and Heidelberg: Lambert Schneider Verlag.
- Buber, M. (1952). *Hope for this hour*. In Buber, M. (1957) *Pointing the Way. Collected Essays*. New York: Harper & Brothers.
- Buber, M. (1953). *Good and Evil*. New York: Charles Scribner's Sons.
- Buber, M. (1953a). Gottesfinsternis. In Buber, M. (1962) *Werke. Erster Band: Schriften zur Philosophie [Collected Works. First Volume: Writings in Philosophy]*. München and Heidelberg: Lambert Schneider Verlag.
- Buber, M. (1954). Elemente des Zwischenmenschlichen. In Buber, M. (2006). *Das dialogische Prinzip*. Gütersloh: Lambert Schneider/Gütersloher Verlagshaus.
- Buber, M. (1957). Schuld und Schuldgefühle [Guilt and Feelings of Guilt]. In Buber,

- M. (1962) Werke. Erster Band: Schriften zur Philosophie [Collected Works. First Volume: Writings in Philosophy]. München and Heidelberg: Lambert Schneider Verlag.
- Buber, M. (1957a). Nachwort: Zur Geschichte des dialogischen Prinzips [Afterword. On the History of the dialogical Principle]. In Buber, M. (2006). *Das dialogische Prinzip [The Dialogical Principle]*. Gütersloh: Lambert Schneider/Gütersloher Verlagshaus.
- Buber, M. (1958). *I and Thou*. Edinburgh: T. & T. Clark.
- Buber, M. (1961) *Aus einer Philosophischen Rechenschaft*. In Buber, M. (1962) Werke. Erster Band: Schriften zur Philosophie [Collected Works. First Volume: Writings in Philosophy]. München and Heidelberg: Lambert Schneider Verlag.
- Buber, M. (1963) Antwort [Response]. In Schilpp, P.A. & Friedman, M. (Eds.) *Martin Buber*. Stuttgart: W. Kohlhammer Verlag.
- Buber, M. (1964). Philosophical Interrogations. In Rome, S. & Rome, B. *Philosophical Interrogations*. New York: Holt, Rinehart and Winston.
- Buber, M. (1970) *I and Thou*. New York: Charles Scribner's Sons.
- Buber, M. (1986). *Begegnung. Autobiographische Fragmente [Encounter. Autobiographical Fragments]*. Heidelberg: Verlag Lambert Schneider.
- Buber, M. (1995). *Ich und Du [I and Thou]*. Stuttgart: Philipp Reclam jun. GmbH & Co.
- Buber, M. (1996). *The Way of Man: According to the Teachings of the Hasidim*. New York: Kensington Publishing Corporation.
- Buber, M. (2002). *Between Man and Man*. London: Routledge Classics.
- Buber, M. (2004). *I and Thou*. London: Continuum.
- Chisholm, R. M. (1999). The Persistence of Persons. In Kim, J. & Sosa, E. (Eds.), *Metaphysics. An Anthology*. Malden: Blackwell Publishers.
- Collins, S. (1982). *Selfless Persons. Imagery and Thought in Theravada Buddhism*. Cambridge: Cambridge University Press.
- Damasio, A. (2000). *The Feeling of What Happens*. London: Vintage.
- Dennett, D. (1992). The Self as a Center of Narrative Gravity. In Kessel, F., Cole, P. & Johnson, D. (Eds.). *Self and Consciousness: Multiple Perspectives*. Hillsdale: Lawrence Erlbaum Associates.
- Drengson, A. (2008). The Life and Work of Arne Naess: An Appreciative Overview.

- In Drengson, A. & Devall, B. (Eds.). *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Doris, J. M. (2002). *Lack of Character: Personality and Moral Behavior*. New York: Cambridge University Press.
- Elster, J. (1985). *The Multiple Self (Studies in Rationality and Social Change)*. Cambridge: Cambridge University Press and Universitetsforlaget AS (Norwegian University Press).
- Fine, C. (2005). *A Mind of Its Own. How your Brain Distorts and Deceives*. Cambridge: Icon Books.
- Feinberg, T. E. (2001). *Altered Egos. How the Brain Creates the Self*. Oxford: Oxford University Press.
- Flanagan, O. J. (1991). *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, Massachusetts: Harvard University Press.
- Flanagan, O. J. (2009). Moral Science? Still Metaphysical After All These Years. In Narvaez, D. & Lapsley, D. K. (2009). *Moral Personality, Identity and Character: Explorations in Moral Psychology*. Cambridge University Press.
- Fox, M. (1963). Einige Probleme in Bubers Moralphilosophie [Some Problems in Buber's Moral Philosophy]. In Schilpp, P. A. & Friedman, M. (Eds.) *Martin Buber*. Stuttgart: W.Kohlhammer Verlag.
- Frankfurt, H. G. (1998). Freedom of the Will and the Concept of a Person. In Frankfurt, H. G. *The Importance of What We Care About. Philosophical Essays*. Cambridge: Cambridge University Press.
- Fraser, N. & Honneth, A. (2003). *Redistribution or Recognition? A Political-Philosophical Exchange*. London and New York: Verso.
- Friedman, M. (1963). Die Grundlagen von Martin Bubers Ethik [The Foundations of Martin Buber's Ethics]. In Schilpp, P. A. & Friedman, M. (Eds.) *Martin Buber*. Stuttgart: W.Kohlhammer Verlag.
- Friedman, M. (1981). *Martin Buber's Life and Work. The Early Years 1878-1923*. New York: E.P.Dutton Publishing.
- Friedman, M. (2005). Martin Buber and Mikhail Bakhtin: The Dialogue of Voices and the Word that is Spoken. In: Banathy, B. H. & Jenlink, P. M. (Eds.) *Dialogue as a Means of Collective Communication*. New York: Kluwer Academic Publishers.
- Glover, J. (2001). *Humanity. A Moral History of the Twentieth Century*. New Haven: Yale University Press.
- Goethe, J. W. (1967). *Faust*. Leipzig: Insel-Verlag.

- Hacking, I. (1995). *Rewriting the Soul. Multiple Personality and the Sciences of Memory*. Princeton, New Jersey: Princeton University Press.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*. 108, 814-834
- Haidt, J. (2008). *Social Intuitionists Answer Six Questions about Moral Psychology*. In: Sinnott-Armstrong, W. (Ed.). *Moral Psychology. Volume 2. The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, Massachusetts: MIT Press.
- Harré, R. (1998). *The Singular Self. An Introduction to the Psychology of Personhood*. London: Sage Publications.
- Hobbes, T. (1996). *Leviathan*. Oxford: Oxford University Press.
- Honneth, A. (1994). *Kampf um Anerkennung: Zur Grammatik sozialer Konflikte*. Frankfurt am Main: Suhrkamp Verlag.
- Hornsby, J. (2004). Agency and Actions. In Hyman, J. & Steward, H. (Eds.). *Agency and Action*. Cambridge: Cambridge University Press.
- Hume, D. (1907). *An Enquiry concerning the Principles of Morals*. Chicago: The Open Court Publishing Co.
- Hume, D. (2009). *A Treatise of Human Nature. Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*. The Floating Press.
- Iacoboni, Molnar-Szakacs, Gallese, Buccino, Mazziota et al. (2005). Grasping the Intentions of Others with One's Own Mirror Neuron System. *PLoS Biol* 3(3): e79. doi:10.1371/journal.pbio.0030079
- Kant, I. (1877). *Kritik der Reinen Vernunft*. Wiesbaden: VMA-Verlag.
- Kant, I. (1974). *Kritik der praktischen Vernunft*. Frankfurt am Main: Suhrkamp Taschenbuch Verlag.
- Kant, I. (1991). *The Metaphysics of Morals*. Cambridge: Cambridge University Press.
- Kant, I. (1996). Was ist Aufklärung? [What is Enlightenment?] In Bahr, E. (Ed.). *Was ist Aufklärung? Thesen und Definitionen*. Stuttgart: Philipp Reclam jun.
- Kant, I. (1999). *Grundlegung zur Metaphysik der Sitten*. Hamburg: Meiner Verlag.
- Kaufman, W. (1996). Acknowledgments. In Buber, M. (1996). *I and Thou*. New York: Touchstone.
- Kennett, J. and Fine, C. (2008). Internalism and the Evidence from Psychopaths and

- ‘Acquired Sociopaths’. In Sinnott-Armstrong, W. (Ed.). *Moral Psychology. Volume 3. The Neuroscience of Morality: Emotion, Brain Disorders, and Development*. Cambridge, Massachusetts: MIT Press.
- Kerr, F. (2004). The Self and the Good: Taylor’s Moral Ontology. In Abbey, R. (Ed.): *Charles Taylor*. Cambridge: Cambridge University Press.
- Keynes, M. J. (1972). My Early Beliefs. In *The Collected Writings of John Maynard Keynes. Volume X. Essays in Biography*. London: Macmillan St. Martin’s Press for the Royal Economic Society.
- Krone, W. (1993). *Martin Buber – Erziehung unter dem Radikalanspruch mitmenschlicher Verantwortung. Überlegungen zur Verantwortungsproblematik im Spätwerk Martin Bubers aus pädagogischer Sicht*. Frankfurt am Main: Peter Lang Verlag.
- Korsgaard, C. (2009). *Self-Constitution. Agency, Identity and Integrity*. Oxford: Oxford University Press.
- Kymlicka, W. (2002). *Contemporary Political Philosophy. An Introduction*. Oxford: Oxford University Press.
- La Fintaine, J.S. (1985). Person and Individual: Some anthropological reflections. In M. Carrithers, S. Collins, & S. Lukes, *The Category of the Person*. Cambridge: Cambridge University Press.
- Levinas, E. (1963). Martin Buber und die Erkenntnistheorie. In P. A. Schilpp, P.A. & M. Friedman (Eds.) *Martin Buber*. Stuttgart: W. Kohlhammer Verlag.
- Lindblom, J. & Ziemke, T. (2002). Social Situatedness: Vygotsky and Beyond. In Prince, Demiris, Marom, Kozim & Balkenius (Eds.) *Proceedings of the Second International Workshop on Epigenetic Robotics: Modelling Cognitive Development in Robotic Systems* (pp. 71-78). Lund University Cognitive Studies, Vol. 94. <http://cogprints.org/2517/1/Lindblom.pdf>
- Martin, R. & Barresi, J. (2003). Introduction: Personal Identity and What Matters in Survival: An Historical Overview. In R. Martin & J. Barresi (Eds.). *Personal Identity*. Oxford: Blackwell Publishing.
- Mauss, M. (1985). A Category of the Human Mind: The Notion of Person; the Notion of Self. In M. Carrithers, S. Collins & S. Lukes (Eds.) *The Category of the Person*. Cambridge: Cambridge University Press.
- McGeer, V. (2008). Varieties of Moral Agency: Lessons from Autism (and Psychopathy). In: W. Sinnott-Armstrong (Ed.), *Moral Psychology. Volume 3. The Neuroscience of Morality: Emotion, Brain Disorders, and Development*. Cambridge, Massachusetts: MIT Press.
- Mendez, M. F. & Gershfield, D. N (2004). Delirium. In: W. G. Bradley, R. B.

- Daroff, G.M. Fenichel & J. Jankovic (Eds). *Neurology in Clinical Practice. Principles of Diagnosis and Management*. Philadelphia: Butterworth Heinemann.
- Mill, J. S. (1987). Utilitarianism. In: J. S. Mill & J. Bentham (1987). *Utilitarianism and Other Essays*. London: Penguin Books.
- Mill, J. S. (2002). *The Basic Writings of John Stuart Mill. On Liberty, The Subjection of Women & Utilitarianism*. New York: Modern Library.
- Naess, A. (1973). The Place of Joy in a World of Fact. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Naess, A. (1982). Spinoza and the Deep Ecology Movement. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Naess, A. (1985). Gestalt Thinking and Buddhism. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Naess, A. (1986). Self-Realization: An Ecological Approach to Being in the World. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Naess, A. (1992). An Example of a Place: Tvergastein. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Naess, A. (1992a). Sustainability! The Integral Approach. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Naess, A. (1993). Beautiful Action: Its Function in the Ecological Crisis. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Naess, A. (1995). Industrial Society, Postmodernity & Sustainability. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Naess, A. (2005). The Basics of the Deep Ecology Movement. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.
- Naess, A. (2006). Pluralism in Cultural Anthropology. In: A. Drengson & B. Devall (Eds.) *Ecology of Wisdom. Writings by Arne Naess*. Berkeley: Counterpoint Press.

- Nietzsche, F. (1996). *On the Genealogy of Morals*. Oxford: Oxford University Press.
- Nagel, T. (2009). The I in Me. On: Selves. An Essay in Revisionary Metaphysics by Galen Strawson. *London Review of Books*. Vol. 31, No. 21. 5 November 2009.
- Olson, E. T. (1998). There is No Problem of the Self. *Journal of Consciousness Studies*, 5, No. 5-6, pp. 645-57
- O'Shea, M. (2005). *The Brain. A Very Short Introduction*. Oxford: Oxford University Press.
- Parfit, D. (1986). *Reasons and Persons*. Oxford: Oxford University Press.
- Pinker, S. (2002). *The Blank Slate. The Modern Denial of Human Nature*. London: Penguin Books.
- Putnam, H. (2008). *Jewish Philosophy as a guide to life. Rosenzweig, Buber, Levinas, Wittgenstein*. Bloomington and Indianapolis: Indiana University Press.
- Ramachandran, V.S. (2006, October). Mirror Neurons and the Brain in the Vat. *Edge*. Retrieved from http://www.edge.org/3rd_culture/ramachandran06
- Ramachandran, V.S. (2000, June). Mirror Neurons and Imitation Learning as the Driving Force behind "The Great Leap Forward" in Human Evolution, *Edge*. Retrieved from http://www.edge.org/3rd_culture/ramachandran/ramachandran_p1.html
- Reichert, T. (1996). *Buber für Atheisten. Ausgewählte Texte [Buber for Atheists. Selected Texts]*. Gerlingen: Lambert Schneider Verlag.
- Regan, T. (1985). *The Case for Animal Rights*. Berkeley: University of California Press.
- Russell, B. (1952). Is There a God?. In J. C. Slater & P. Kollner (Eds.). *The Collected Papers of Bertrand Russell, Volume 11: Last Philosophical Testament*. London: Routledge.
- Sabini, J. & Silver, M. (1982). *Moralities of Everyday Life*. Oxford: Oxford University Press.
- Searle, J. (1984). *Minds, Brains and Science*. Cambridge, Massachusetts: Harvard University Press
- Seigel, J. (2005). *The Idea of the Self. Thought and Experience in Western Europe since the Seventeenth Century*. Cambridge: Cambridge University Press.
- Simmel, G. (1906). *Zur Soziologie der Armut*. In E. Jaffé, W. Sombart & M. Weber

(Eds.) *Archiv für Sozialwissenschaft und Sozialpolitik*. 22.Jg. (N.F. 4), 1. Heft (January). Retrieved from: <http://socio.ch/sim/arm06.htm>

- Simon, E. (1963). Martin Buber, der Erzieher. In: Schilpp, P.A. & Friedman, M. (Eds.) *Martin Buber*. Stuttgart: W. Kohlhammer Verlag.
- Singer, P. (1986). All Animals are Equal. In P. Singer (Ed.). *Applied Ethics*. Oxford: Oxford University Press.
- Shoemaker, S. (1970). Persons and their Pasts. In Kim, J. & Sosa, E. (Eds.), *Metaphysics. An Anthology*. Malden: Blackwell Publishers.
- Shweder, R. A. & Haidt, J. (1993). The Future of Moral Psychology: Truth, Intuition, and the Pluralist Way. *Psychological*. Vol. 4, No. 6.
- Smith, R. G. (1967). *Martin Buber*. Richmond, Virginia: John Knox Press.
- Sorabji, R. (1993). *Animal Minds and Human Morals: The Origins of the Western Debate*. New York: Cornell University Press.
- Sorabji, R. (2006). *Self. Ancient and Modern Insights about Individuality, Life and Death*. Oxford: Oxford University Press.
- Sperber, D. (2005). Modularity and Relevance. How can a Massively Modular Mind be Flexible and Context-Sensitive? P. Carruthers, S. Laurence / S. Stich (Eds.), *The Innate Mind. Structure and Contents*. Oxford: Oxford University Press.
- Strawson, G. (1999). The Self. In S. Gallagher & J. Shear (Eds.). *Models of the Self*. Exeter: Imprint Academic.
- Taylor, C. (1985). *Human Agency and Language. Philosophical Papers I*. Cambridge: Cambridge University Press.
- Taylor, J. B. (2008). *My Stroke of Insight*. London: Hodder & Stoughton.
- Thomson, J. J. (1977). *Acts and Other Events*. Ithaca and London: Cornell University Press.
- Varela, F. J. (1999). *Ethical Know-How. Action, Wisdom and Cognition*. California: Stanford University Press.
- Wilkes, K. V. (1999). ΓΝΩΘΙ ΣΕΑΥΤΟΝ (Know Thyself). In S. Gallagher & J. Shear (Eds.). *Models of the Self*. Exeter: Imprint Academic.
- Wise, S. M. (2000). *Rattling the Cage. Towards Legal Rights for Animals*. Cambridge, Massachusetts: Perseus Publishing.