# THE PHENOMENON OF AMORALISM:
## An Investigation of the Cognitive and Emotive Roots

by

ANDREI G. ZAVALIY

A dissertation submitted to the Graduate Faculty
in Philosophy in partial fulfillment of the requirements
for the degree of Doctor of Philosophy,
The City University of New York

2008

UMI Number: 3296921





ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

Date: 10/17/2007 Steven Cahn

Chair of Examining Committee

Date: 10/17/2007 John Greenwood

Executive Officer

Supervisory Committee:

Professor Douglas Lackey _____

Professor Steven Ross _____

Professor Peter Simpson _____

THE CITY UNIVERSITY OF NEW YORK

THE CITY UNIVERSITY OF NEW YORK

Abstract

# THE PHENOMENON OF AMORALISM:
## An Investigation of the Cognitive and Emotive Roots

by

Andrei G. Zavaliy

Adviser: Professor Douglas P. Lackey

An amoralist is defined as a person who rejects the claims of moral reasons to special authority, and systematically acts without regard to the generally accepted moral standards. A psychopath can be seen as a paradigm case of an *extreme* amoralist, although the less severe cases of *selective* amoralists are considered. The research into the typical behavioral pattern, motivational structure, and the value system of psychopaths can shed light on at least three aspects related to the analysis of the moral agency. First, it can help elucidating the emotive and cognitive conditions necessary for moral performance. Secondly, it can provide empirical evidence supporting the externalist theories of moral motivation. Finally, it can bring into greater focus our intuitive notion of the limits of moral responsibility. In this work I concentrate on the first and the last aspects, but the discussion has an indirect bearing on the second theme as well.

The phenomenon of amoralism presents a challenge: the psychopaths are *not* usually diagnosed as psychotic, and yet there is a sense in which their condition is clearly abnormal and needs explaining. The main purpose of this work is to specify the typical features of

amoral individuals and point to the roots of this abnormality. The two explanatory options are considered: the amoralist may suffer either from emotive deficiency, or else exhibit a cognitive failure of some sort. The comparison of amoralists with autistic individuals allows dismissing the emotive deficiency as the main reason of the amoral condition, and an argument is developed to show that a fundamental mistake in judgments about good and evil is ultimately responsible for this condition.

The amoralists, however, can be held morally responsible for their actions. I seek to show that certain arguments that encourage us to see the psychopath as a badly wired machine that is exempt from responsibility are wrong. Finally, it is argued that the libertarian interpretation of a free choice puts an important constraint on our ability to explain the basis of one's decision to adopt a set of values incompatible with the essential values of the moral point of view.

For my wife Olga,

*sine qua non.*

# ACKNOWLEDGMENTS

It has been a long, tiring but enormously rewarding intellectual journey since the time I laid my hands on the very first book in philosophy (a *Short Dictionary of Philosophy* that I accidentally discovered among my father's books at the age of thirteen), and the time when I completed the doctoral degree in the same discipline. On more than one occasion that particular event in my early adolescence has given me food for thought about the exact role of chance in determination of one's professional calling.

Needless to say, the journey itself and the present work in particular would not be possible without the help and generous support of many people along the way. First of all, I am greatly indebted to the two philosophy professors who supported me in many ways and always challenged me intellectually during my undergraduate studies at Nyack College: Dr. Ronald Ruegsegger and Dr. James Danaher. I want to thank Dr. Ruegsegger for teaching me through his writings and his inimitable lecturing style the value of systematic method in philosophy and for being an example of a responsible and thoroughgoing scholar. I am thankful to Dr. Danaher for trusting in me, for constantly encouraging and inspiring me in my philosophical wanderings throughout the college years, and for being a true role model, not as a philosopher only, but, above all, as a fulfilled human being.

The list of the professors at the Graduate School who in one way or another contributed to my philosophical development, and are thus partially responsible for the present work, would be long indeed. But at least several names should certainly be mentioned here. I want to thank Prof. Peter Simpson for his thorough reading of the earlier versions of this text, many valuable suggestions, substantive disagreements with my own

position, but, more importantly, for his steadfast belief in *natural goodness*, which was always very comforting during my immersion into the world of amoralists.

Much is owed in this work to the discussions I had with Prof. Steven Ross, both in his office at Hunter College and over the internet. His distinctive, versatile teaching style, his unceasing enthusiasm for the subject matter of moral philosophy combined with deep respect for students and their philosophical efforts are the qualities I would like to approximate in my own teaching career. More specifically, I want to thank Prof. Ross for showing me how to be a *real* moral realist, without yet falling into the extravagancies of classic intuitionism or the rigidity of reductive moral naturalism. His argument for the non-reductive autonomy of moral terms and moral emotions serves as the necessary theoretical basis for my treatment of amoralism as an abnormal condition.

My main supervisor, Prof. Douglas Lackey, is a person who in many respects introduced me to the unfamiliar world of graduate studies, being one of the first philosophy professors I encountered after leaving the comfort zone of college. His series of lectures on Hegel will remain for me as an example of how to present the most difficult material in a way that is both accessible and yet faithful to the original source. Throughout the 'active phase' of writing of dissertation his many non-intrusive, almost tacit suggestions have significantly altered my vision of the problem, and, as a result, the course of the main argument, and in so many ways have helped me to make a transition from the stage of the first Dissertation Prospectus, still very much clouded in ambiguity and indecisiveness, to the final version below, completed almost two years later.

A special mentioning goes to Prof. John Greenwood, the Executive officer of the Philosophy Program, whose favorable reading of my work helped to secure an acceptance for

publication of the part of this dissertation dealing with the moral responsibility of the psychopaths. This is besides him being helpful and responsive in so many other ways throughout my years at the Graduate School.

Finally, I would like to give special thanks to my family, and, above all, to my wife Olga (and, of course, to my now two-year old son Alexander who was born approximately in the period between the first and the second chapters of this work), for believing in me and supporting me during all these years. Many of the ideas that eventually were incorporated in this work (and many more that were rejected) were initially tested against her superb common sense, uncontaminated by the philosophical dogmas and ideological agendas. Having such a critic close at hand proved to be extremely valuable in the long run. On the whole, it is her patient love and care that enabled me to complete this work, and it is to her that this work is dedicated.

# TABLE OF CONTENTS

# CHAPTER I

## INTRODUCTION AND PRELIMINARIES

> *Caligula*: So, I take it, you believe in some higher principle.
> *Cherea*: Certainly I believe that some actions are – shall I say? – more praiseworthy than others.
> *Caligula*: And *I* believe that all are on an equal footing.
>
> Albert Camus, *Caligula*.

In this chapter I shall identify the main objective of the project as the search for the explanation of amoralism. The concept of amoralism will be compared and contrasted with a number of related concepts (e.g., immoralist, moralist), and an acceptable definition of amoralism will be worked out and defended against several alternative conceptualizations of this phenomenon. I shall argue that being *amoral* is a matter of degree, and that it is useful to distinguish at least two variations of this general condition. Finally, I shall provide an outline of all the subsequent chapters of this dissertation.

### 1. The Statement of the Problem

It is a common empirical observation that most people take considerations of morality as important factors in making a morally relevant choice. Moreover, such considerations ordinarily override other reasons for actions. To say that a certain act is useful, agreeable, profitable but yet immoral usually means to condemn the act in the strongest possible terms. In the ordinary discourse the labeling of some action-guiding principle or a particular action as immoral gives the ultimate reason for rejecting it. Likewise, one may think of an action as difficult, unpleasant, risky and costly but yet required from the moral point of view. Pointing

to moral obligation in this case, too, functions as the reason that (normally) outweighs any other consideration.[1]

The normative overridingness of moral principles[2] does not imply that people on all occasions choose to act morally, i.e., it does not imply *motivational* overridingness. The requirements and prohibitions of morality can often be overshadowed in the actual life situation by reasons of self-interest, or ignored due to inattentiveness, weakness of will and the like. The somewhat fragile status of moral considerations in certain real life decisions does not yet undermine the legitimacy of the claims of morality to the highest authority.[3] The observation that most people under normal conditions respect the limits set by the moral point of view and act accordingly, bears witness to the claim that the authority of morality is not merely announced but is also widely recognized by the competent participants in social life.

But there exists a class of human beings who conspicuously differ in this respect from the majority. These persons stand out because they take no notice of the community's accepted standards of right and wrong, and remain unmoved by the ordinary practical moral reasoning in general. Moreover, if we are to relegate them to a class of their own, such a disregard for morality on the part of these individuals must be much more thorough and more manifest

---

[1] At this early point, the claim about the priority and overridingness of moral reasons is not to be taken as an uncritical assumption of the strong Kantian view, according to which moral judgments are derived from reason and "when properly grounded, express its full legislative authority" (Audi, R., 2006, p. 159). I am offering this claim as a commonsensical observation about the ordinary cases of moral reasoning. Whether one explains moral values as deriving from God, reason or society, or even has no theoretical justification of morality at all, taking moral reasons as *authoritative* seems to be a product of any 'civilized' upbringing.

[2] I understand by a moral principle either a very general moral statement that claims to specify the conditions under which an action is right or wrong, or a less general command that states that some specific type of action is morally permissible, obligatory, or prohibited. Thus defined, I shall be using 'moral principle' and 'moral rule' interchangeably.

[3] I am making a relatively uncontroversial claim here that pointing to the cases of moral failure or noncompliance with moral requirements is not by itself sufficient to establish a universal moral skepticism or nihilism.

than the commonly recognized instances of the occasional moral failures in the otherwise upright members of society. There must be present either an explicit contempt for the moral reasons or systematic indifference to all moral considerations which goes beyond the more familiar cases of immoral actions due to intellectual carelessness, distraction and inattentiveness to the relevant facts.

But before we can look into the concrete cases of such an overt disregard for morality, it is useful first to provide a purely *functional description* of a kind of individual who will be the main object of analysis in this work.[4]

He[5] is a person whose specific actions, outspoken beliefs, and the whole way of life show no regard whatsoever for the moral dimension of the social life. Moral discourse is often an unknown foreign language to him. Appeals to categories of rightness and wrongness make no noticeable influence on his behavior, and, in some extreme cases, he might be genuinely puzzled by the notions of guilt, moral blame and responsibility. The predicament and intensity of a moral conflict escapes his understanding; the concept of retributive justice makes no sense to him. He remains unmoved by the sight of suffering of other people. The perceived self-interest is usually the only motive that drives his behavior, and the thought about the interests of others around him plays no role in making his everyday decisions. His actions are often impulsive and unpredictable, and when prompted to provide an explanation of his particular choices, he rarely goes beyond the notorious "because I wanted to."

A person who fits (at least partially) the above description is surely not an unknown breed to moral philosophy and classical literature, nor to actual human history. Callicles in

[4] This description will necessarily be of an ideal or *archetypal* amoralist, and may not apply in all details to real cases that will be discussed later.

[5] Throughout, and merely for convenience, I use the masculine pronoun in referring to amoralist.

Plato's "Gorgias" and Thrasymachus in "The Republic", Hobbes' '*Foole'*, Hume's 'Sensible Knave', Mandeville in his *Fable of the Bees*, Hegel's 'World-historical individual' and Nietzsche's 'Übermensch' might all be interpreted (to a greater or lesser degree) as paradigmatic instances of the general type. On the more mundane side, one might mention sociopaths, persons suffering from narcissism and certain (but not all) autistic individuals as exemplifying some of the traits mentioned above. What all these characters seemingly have in common is the fact that they, willingly or unwillingly, position themselves (at least temporarily) outside the existing system of moral rules and requirements, and thus refuse to accept the ordinary evaluative moral pronouncements about this type of action or this type of a life-style as applicable to them.[6]

To give such an individual a convenient label, and following the established tradition, I shall call him an *amoralist*, and his general position (or, perhaps, his condition) with regard to morality will be referred to as *amoralism,*[7] even if the name will require further qualifications as we go on.

Despite the fact that the amoralist is frequently mentioned in literature as a philosophically significant character, it appears that in most cases the authors who discuss the issue tend to have a predominantly *instrumental* interest in this type.[8] The fact that the

---

[6] Whether they ever succeed in 'stepping outside' the realm of morality (i.e., whether it is indeed a possible stance), and whether one can escape the force of a moral judgment simply by refusing to accept its legitimacy, is, of course, a question that will have to be addressed later. Bernard Williams, for instance, expresses a long-standing tradition which denies such possibility when he says: "The moral law is more exigent than the law of an actual liberal republic, because it allows no emigration." (1985, p. 178).

[7] 'Extreme moral nihilism' is an alternative term used, among others, by G. Harman, who describes it as follows: "An extreme version of nihilism holds that morality is simply an illusion: nothing is ever right or wrong, just or unjust, good or bad." (1977, p. 11).

[8] This is not to deny the legitimacy and fruitfulness of this instrumental use of the notion of amoralism. Ronald Milo (1984) seems to be a notable exception to this rule, although his major concern is still with the analysis of the *concept* of immoralism rather than with searching for the roots of this phenomenon.

possibility of a genuine amoralist is apparently quite coherent is often seen as a proof of *externalism* in moral motivation,[9] that is, as a strong evidence for the 'division of labor' between beliefs or pure cognitive states, which are states representing the way the world is, and the conative states (e.g., desires) which alone can bring about an intentional action. In other cases, the amoralist is introduced as a theoretical opponent, or even as an imaginary 'boogeyman', against whom the moral point of view must be grounded or rationally justified, and whose function in those cases is much like the function of a universal skeptic from Descartes' "First Meditation". Once the morality is thus grounded it is then presumably the task of the moral philosopher to discover some cognitive failure or logical inconsistency in the position of an amoralist. Still other authors discuss the amoralist only to deny his actual existence as a possible, let alone attractive, way of life. Once the position of a thoroughgoing amoralist is fully spelled out, it is argued, it will become apparent that it is now an abnormal medical condition rather than an alternative intellectual stand or a defensible life policy.[10]

In the present work, however, I shall not be using the fact of amoralism merely as a theoretical tool to establish some further agenda. It will be argued that the philosophical significance of amoralism for moral theory is not exhausted by the limited role usually assigned to it. My primary interest lies in the phenomenon of amoralism as such, apart from the question whether it can be used to support or weaken any particular position in ethics. I want to suggest that by itself it is a *puzzling* phenomenon, i.e., the one that requires some sort of explanation. Elucidating the underlying features of the actual cases of amoralism, discovering its likely causes, such as cognitive and emotive deficiencies, as well as considering the degree of responsibility of the amoralists for their actions, is a project worthy

[9] E.g., Brink (1997).

[10] E. g., Williams (1972).

of pursuance in its own right. The prospect of the greater understanding of the *moral* agency in the process of the analysis of the *amoral* one gives an additional impetus to this work.

The main objective of this work is to a great degree *explanatory*, and as such, it will remain neutral with regard to a number of questions in moral theory that might naturally arise in the course of the discussion. At the same time, a successful research project of this kind might perhaps *suggest* the definite answers to a number of meta-ethical questions, and lend some *indirect* support to certain views in moral philosophy such as externalism in moral motivation, externalism about normative reasons, and cognitivism in meta-ethics.[11] This last result, however, should be seen as a side effect rather than the main aim of the present work.

## 2. Conceptual Background: Amoralism and Related Notions

We must now go beyond the initial functional description of the amoralist given in the previous section and try to capture the specific behavioral and mental characteristics of such a person by means of a formal definition. One helpful approach here is to start with the demarcation of the notion of amoralism from the several kindred categories. The most important ones for our purposes are the notions of *moralism* and *immoralism*. The difference between these concepts must be clearly stated before we can finally focus on the amoralist.

Richard Garner is one of the few contemporary philosophers who tried to formally define the amoralist and contrast his point of view with both that of a *moralist* and an *immoralist*. The notion of a *moralist* is the more basic and simpler notion in this context, and thus should be addressed first. Garner claims that "We qualify as 'moralists' if we believe the

[11] For the claim that moral *cognitivism* and motivational *externalism* are better equipped for explaining the possibility of amoralism see Brink, 1997 and Devitt, 2002.

world contains values to discover, rules to follow, or rights to recognize."[12] Moreover, a moralist is the one who takes morality seriously, and tries to regulate his behavior in accordance with the requirements and constraints of the existing moral rules. Being a moralist, on this view, is a matter of having certain beliefs as well as acting in a certain way. This minimalist definition is general enough to allow for almost any interpretation of the ultimate nature of morality, its foundations and justification. Whether the moral code in question is God-sanctioned, reason-sanctioned or is a matter of the conventional regulations that spontaneously appeared in the course of the prolonged communal living, it would not affect the acceptability of the above definition. In this sense, there clearly *are* moralists,[13] and even if a specific psychological or social explanation of their behavior might be contested, their existence cannot.

The opposite of a *moralist* is somewhat more difficult to specify. One can fail to confirm to morality in a variety of ways and for a variety of reasons, and it might seem that every specific occasion of a moral failure on the part of a moral agent can be generalized into a class of its own, with its own distinctive features. A person who breaks the relevant moral principles by the reason of insanity, or the one who systematically commits wrongful acts by accident and neglect, as well as the calculative thief who violates the property rights of another with purely selfish motives, and a self-proclaimed 'higher being' who allegedly stands above the 'merely human' morality, are all *non-moralists*, but it is clearly unhelpful for our purposes to place all of them into the same category.

The distinction that is frequently drawn when discussing the non-moral types is that between an *immoralist* and an *amoralist*. Since there is a considerable degree of vagueness in

[12] Garner, R., 1994, p.1.
[13] Another common meaning of the word 'moralist' as referring to a person unduly concerned with the morals of others will be irrelevant here.

the ordinary speech as to when the one or the other notion applies, the précising definition of both terms will be, to a large extent, stipulative, and the suitable criterion that would allow us to distinguish between the two should be determined in view of the overall objective of this work. Once again, Richard Garner offers a convenient starting point:

> People who do not *follow* the ethical and moral requirements of the society to which they belong can be called *immoral* (at least by the moralists in their society), but a person who is *amoral*, an amoralist, does not believe certain claims made about those requirements, however he or she behaves. The moral standards we accept (if any) determine who (if anyone) we call immoral, but everyone should call an amoralist an amoralist.[14]

On Garner's view, the crucial difference between an amoralist and immoralist lies not in their respective actions, but rather in their beliefs about the nature of morality. On the one hand, the immoralist is recognized by his (immoral) actions, but it is conceivable that the behavior of the amoralist confirms to the accepted moral standards in a given society, and what makes him an amoral (rather than an immoral) person is the lack of certain essential beliefs about moral principles and moral values, which, presumably, the immoralist still retains despite his behavioral failures.[15] The amoralist may even be a competent participant in moral discourse:

> It is possible to "reject morality" without rejecting the language of morality, and possible for those who "abandon" morality to act in ways people would describe as morally admirable. One's amoralism does not reside in what one says, or even how one acts. It is a matter of *what one believes* about morality.[16]

---

[14] Garner, R., 1994, p.17.

[15] We may also observe that for Garner the extension of 'immoral' varies with the culture, while the extension of 'amoral' is fixed. The moral standards of a particular society will determine the set of all immoralists (identified primarily by their behavior), but not the set of all amoralists.

[16] *Ibid*., p. 43.

Those essential beliefs which, according to Garner, are absolutely necessary for remaining 'inside' morality include, among others, beliefs "about the objectivity and prescriptivity of rules, prohibitions, virtues, vices, rights, and duties."[17] In addition, for the amoralist, all moral goodness and moral obligations are "of the relative kind", nothing is "good in itself", and no person for him has "intrinsic worth".[18]

I see two weaknesses with Garner's classification. First, Garner's emphasis on the lack of certain *beliefs* as the distinguishing mark of the amoralist, together with the implied non-behaviorist account of the nature of beliefs in general,[19] warrant the conclusion that it is quite *possible* for an intelligent amoralist to remain undetected as long as his external behavior does not (significantly) deviate from the accepted standards. Even the occasional moral lapses on the part of the amoralist would not reveal his true identity, since these can always be attributed to weakness of will, and the like. Garner's definition implies, as he admits, that "there are moralists who lie, cheat, steal, and despise humanity; and there are amoralists who tell the truth, play fair, and treat others with respect."[20] As a result, the extent of the *amoralism* in the society might be, for all we know, much greater than anyone ever suspected, and the cases of genuine amoralism (which, *ex hypothese*, can never be detected empirically[21]) loose much of their significance for ethical theory.

---

[17] *Ibid*., p. 17.

[18] *Ibid*., p. 19.

[19] This seems to follow from his claim that one can conform to accepted morality, i.e., behave in a certain way, without believing that one (morally) ought to do so.

[20] *Ibid*., p. 279. This may sound less offensive if we keep in mind Garner's main thesis, namely, that belief in objective morality is not at all necessary for a good life. Various non-moral constraints can be as successful in fulfilling the function usually taken by morality.

[21] Garner agrees that "it is not easy to judge how much positive effect a belief in objectively binding moral principles will have on a person's behavior." (p. 279).

Furthermore, Garner's definition of amoralism seems to be wide enough to place anyone who has doubts about the 'objectivity' of moral rules, or who is unsure whether there are any non-relative moral obligations, or whether anything is 'good in itself' in the same category with the serial killers and extreme sociopaths. This looks paradoxical at least. A person who questions the objective, non-relative status of the moral prescriptions is usually called a moral relativist or a moral skeptic. The notion of a moral skeptic is an important notion for moral philosophy in its own right, and Garner has not given us good reasons why we should conflate it with that of an amoralist. It might well be the case that all amoralists are moral skeptics and relativists, but the reverse does not always hold. The moral relativists and skeptics can sincerely accept the existing moral order, and act accordingly, without, at the same time, retaining any illusions about its objectivity, rationality, absolute justification, and the like.[22] In this case, we should not say that the skeptic 'rejects morality' altogether and thus qualifies as an amoralist.

While Garner's definition of amoralism might be too broad, the description of an amoralist offered by Bernard Williams appears to make this category *too* exclusive. Williams' reasoning takes the form of a dilemma. If, on the one hand, we consistently strip the amoralist of *all* the relevant characteristics of a moralist (e.g., having any concern for others, being affected by distress of others, etc.) then "this raises the question, whether we have left him enough to be [recognizably human]."[23] The complete negative specification of

[22] Cf.: "To be skeptical about ethics is to be skeptical about the force of ethical considerations. […] But we should not assume that the skeptic must be someone who leads a life that goes against ethical considerations. Perhaps we should rather say that he leaves room for such a life." (Williams, B., 1985, p. 25.)

[23] Williams, B., 1972, p.10.

an amoralist would only produce a psychopath, and, as a result, undermine the traditional alleged importance of amoralism as a genuine threat to morality:

> The amoralist seemed important because he seemed to provide an alternative; his life, after all, seemed to have its attractions. The psychopath is, in a certain way, important to moral thought; but his importance lies in the fact that he appalls us. His importance does not lie in his having an appeal as an alternative form of life.[24]

But, on the other hand, if we take a less radical approach and leave the amoralist with at least *some* features of a moral person, we have no reason then, according to Williams, to place such an individual in a separate category: "If we grant a man with even a minimal concern for others, then we do not have to ascribe to him any fundamentally new kind of thought or experience to include him in the world of morality."[25] A person who is poorly equipped with the sympathetic affections necessary for a full-blown moral life, or someone who applies them selectively to the people he happens to care about at the moment, may perhaps be properly called *immoral* on the occasion, but there is no need yet, Williams would argue, to place him 'beyond morality'.

We may notice first, that unlike Garner, who places the main emphasis on having (or lacking) certain *beliefs* about morality, Williams emphasizes the sympathetic concern and affections, i.e., the non-cognitive states of an individual. The way 'into morality' lies through experiencing the world as having intrinsic value, an attitude which in the long run excludes indifference towards others. This has an advantage of allowing the moral skeptic, i.e., the one who doubts (or denies) the possibility of the (final) justification of the existing moral order, to be a full participant in the moral life of a community. The disposition to *care* is logically

---

[24] *Ibid*., pp. 10-11.

[25] *Ibid*., p. 12.

independent from one's beliefs about the ultimate nature of the object of care, or about the alleged prescriptive status of one's attitudes. Furthermore, Williams' classification duly excludes from the realm of morality those exceptional intelligent psychopaths who might have all the necessary beliefs about morality (i.e., necessary on *Garner's* account), without genuinely being part of it. Benjamin Wolman, for instance, in his influential study of the sociopathic personality observes that "bright psychopaths are capable of pretending to feel guilty and profess remorse. They often talk of morality, but they *feel none*. Usually their words mean nothing and do not prevent their actions."[26] On one plausible interpretation of this passage, the psychopaths who are able "to talk of morality" might have all the relevant beliefs about the subject matter, but perhaps what makes them different is the absence of a certain *moral feeling* and, consequently, the absence of motivation to act, which cannot be simply acquired by adding a new true belief about the nature or content of moral rules and prohibitions.[27] Precisely the lack of this feeling or concern for others is what, for Williams, should be sufficient for identifying an amoralist, and there is no need for further inquiry into his beliefs *about* morality.

As we have seen, Williams' description of a 'thoroughgoing' amoralist as a person wholly devoid of empathy, incapable (even to the *minimal* degree) of compassion and concern for others, makes the concept of 'amoralist' coextensive with that of 'extreme

---

[26] Wolman, B., 1987, p. 95. [my emphasis]

[27] Whether it is indeed possible to hold a belief that one is morally *obligated* to do X without being in the least motivated to do X is a matter of considerable debate between externalists and internalists. The internalist thesis states either that being motivated to act upon a moral belief is part of the very meaning of 'accepting a belief', or that having the appropriate belief is causally or conceptually connected with the corresponding desire to act upon it. The externalists, on the other hand, argue that beliefs about morality, no matter how sincerely held, are motivationally inert (precisely because they are *beliefs*), and an independent desire (or conative state) is needed to entail the typical practical consequences of moral considerations. In my interpretation of the psychopath I am not assuming the truth of externalism; I point out that *if* externalism is true, Williams' definition of amoralism has yet another advantage over Garner's.

psychopath'. The individuals who would satisfy Williams' demanding requirements to get into the class of amoralists are extremely rare in ordinary experience. Nor would any of the classic 'philosophical' amoralists mentioned in the beginning of the chapter (e.g., Thrasymachus, 'Sensible knave', and others) qualify for acceptance for obvious reasons: no matter how unconventional their beliefs and actions might be, they are still very much "recognizably human."[28]

I want to suggest that Williams' account of amoralism is too restrictive for the present purposes since it significantly reduces the space for the philosophically interesting analysis of this phenomenon. One way to make more room for such an analysis is to give a greater extension to the term 'amoralism'. Once we can legitimately use the term to encompass both the extreme cases of the psychopaths or sociopaths and the less severe (and the more familiar) forms of amoral behavior, we can then inquire into the features underlying all cases of amoralism (e.g., cognitive or emotive failures), without thereby incurring on the field of psychiatric science proper.[29]

A classification that initially looks more promising for these purposes is offered by Robert Solomon. In his recent book he suggests the following taxonomy:

> The *moral skeptic* doubts or denies the possibility of justification of morality, but accepts morality nonetheless. A more dangerous character is the *amoralist*, a person

[28] Admittedly, the same would also be true given Garner's definition of amoralism. Garner explicitly argues that Callicles (and Thrasymachus, we might suppose), who is conventionally called an amoralist, is in fact a pseudo-amoralist, i.e., someone with a conventionally *immoral* morality (a *kind* of moralist). (1994, p. 284-5).

[29] It is important to note that, according to the consensus of most scholars, psychopathy or antisocial personality disorder is *not* the same as insanity or psychosis (Samenow, 2004). Prichard observed early on that in the 'morally insane' person there is not "any remarkable disorder or defect of the intellect or knowing and reasoning faculties." (Prichard, 1973, p.16). Thus a psychopath is not a helpless victim of physical or social forces beyond his control but is in some sense in control of his actions and thus responsible for his condition (as I shall argue in Chapter VI), This leaves sufficient room for a philosophical analysis of his choices and actions. The *degree* of one's responsibility in these cases, however, is a highly controversial issue and will be discussed later.

> who acts without regard for generally accepted moral principles, and worse yet is the *immoralist,* who recognizes those principles but conscientiously disobeys them.[30]

Without necessarily endorsing Solomon's 'gradation of badness' here, we may observe that his definition of amoralism, as presented against the background of moral skepticism and immoralism, allows greater flexibility for the discussion and analysis of this phenomenon. On the one hand, Solomon, unlike Garner, recognizes a moral skeptic as a distinct character, and clearly defines him in terms of lacking certain beliefs about the justification of moral principles. The amoralist, on the other hand, is characterized *both* by his actions and his beliefs about morality. While the *immoralist* recognizes the validity of moral requirements but fails to act upon them for reasons other than weakness of will,[31] the amoralist rejects the claims of the moral principles to special authority, and, consequently, these principles play no role in making his everyday choices. Finally, even though Solomon does not explicitly define a *moralist*, we may infer that a moralist, for him, is someone who falls in none of the three categories presented, i.e., a person who believes that morality can be justified, accepts the validity of the generally recognized moral principles, and (non-accidentally) respects the limits set by those principles in the course of making his practical decisions.[32]

---

[30] Solomon, R., 2006, p. 26.

[31] I interpret the phrase "conscientiously disobeys them" in a way that excludes cases of weakness of will. On one common interpretation, in cases of weakness of will or incontinence, a person acts against his better judgment due to the overwhelming 'passion' that somehow clouds the recognition of a particular situation as falling under a general moral rule. It is assumed that the required action would ensue in the absence of such non-cognitive interference. *Conscientious* disobedience to moral requirements, on the other hand, seems to imply the presence of *competing* practical reasoning (e.g., a competing practical syllogism), where the conclusion of such 'immoral' reasoning non-accidentally prevails over the conclusion of the 'moral' argument.

[32] In saying that a moralist 'respects' or 'follows' the moral principles one need not assume the moral agent is being a devout Kantian, whose only motivation is to obey the categorical requirements of morality. The terminology of 'following the rules' is not to be taken as an accurate phenomenological account of actual moral motivation. Indeed, the suggestion that, for instance, one helped a friend in need *because* he wanted to fulfill the relevant moral obligation, might be taken as an offensive remark by the person who thus helped his friend. If

There is still a considerable degree of vagueness in the above classification. It is not clear what would count as acceptable 'justification' of morality, i.e., the possibility of which a moral skeptic would want to question, nor is it immediately obvious in what sense does an immoralist "recognize the moral principles" if he fails to follow them (presumably, on more than one occasion). The latter problem, the problem of 'recognition', is especially troubling. An anthropologist studying the life and culture of a distant tribe may come to discover that a prescriptive rule "Always give precedence to men over women in the distribution of food" is part of the traditional morality of that society. There is a clear sense in which the anthropologist *recognizes* this principle as a moral one. If he systematically disobeys the rule while residing in that society, does he fall under the category of amoralism, immoralism or neither? Would the answer differ depending on whom you ask (e.g., a member from that group as opposed to an outsider)?

Intuitively, one would want to respond that the anthropologist recognizes the moral rule in question merely as a cultural *fact* about that society, not as a universal normative statement which would apply to all moral agents regardless of the cultural origin, and that is what makes the difference. But simply refusing to acknowledge certain existing moral principles as applicable *to me* would not by itself be sufficient to escape the unpleasant label. We *also* want to say that in the case described a person is morally *justified* in *not* applying the rule to himself. This is, presumably, what would distinguish the anthropologist who fails to obey the rules of the tribe from the amoralist proper: whereas both the anthropologist and the amoralist refuse to recognize the validity of certain moral principles, the anthropologist is

---

anything, the person cared about the *friend*, not about the abstract rule or principle. But it is still possible in most cases to describe such a behavior from the *third person* perspective in terms of fulfilling an obligation or following the relevant moral requirement. Mill's distinction between the *rule of action* and the *motive of action* is applicable here. (Mill, J. S., 1998).

*justified* in doing so, while the amoralist is not. This, however, will bring us back to the question of justification of morality and the problem of moral skepticism. It appears now (and will be further argued for below) that the very existence of the amoralist depends on there being not only a 'moral sphere' proper,[33] but a justified (or justifiable) system of morality within this sphere, i.e., a system (or any part thereof) that the amoralist would be *unjustified* to reject.[34]

The immoralist, on the other hand, is said to *recognize* the moral principles *qua* moral principles, and yet he conscientiously disobeys them. Once again, lest we increase the number of the immoralists beyond necessity and make the concept all but empty, the immoralist must recognize those principles in some *other* way as compared to the way an anthropologist from the outside recognizes them as being part of the tribe's moral code. He must acknowledge the validity and special authority of those principles, and also recognize that the requirements and prohibitions of the relevant moral rules straightforwardly apply to him. If the distinction is to be preserved, it must be precisely the kind of recognition which was lacking in the case of a genuine amoralist. Despite this 'stronger' recognition, however, the immoralist "conscientiously" refuses to act upon it. One natural explanation of this apparent inconsistency is the presence of the *competing* non-moral practical reasoning (e.g., the one dealing with the considerations of the narrowly-defined, immediate self-interest), the

[33] One can reject or lack morality only if there is something special about moral rules and obligations that would allow us to distinguish them from the various instances of non-moral practical reasoning. There must be a 'moral sphere' to begin with – amoralists (or immoralists) cannot exist in Hobbes' primordial 'state of nature'. I shall argue in Chapter Two that certain constraints on the content and form of a rule must be respected to qualify as a *moral* rule (e.g., it must be universalizable and deal with human welfare); the *moral* rules cannot be identified *simply* by observing what a person or a given group of people take as action-guiding in their actual behavior.

[34] Of course, the other option is to claim that the anthropologist in the example above *is* an amoralist (just as *all* of us would be from the point of view of some incompatible moral system). If all that is needed is the existence of *some* moral system (whether justified or not), then the rejection of this system would constitute amoralism. But then it is not clear why it would be a *bad* (or at least *unfortunate*) thing to be amoral – a connotation strongly suggested both by the common intuitions and Solomon's definition.

conclusion of which *prevails* over the one of the moral argument.[35] Due to the shift in priorities, the immoralist prefers to act against the prohibitions of morality despite his open acknowledgement of his being subject to its demands.

There is still much that requires further clarification and explanation. The following chapters will address the problem of justification of a moral system and acceptance or recognition of the moral principles in greater details. But even at this early point there emerged a rough conception of amoralism that can be usefully employed in the following analysis, which, in turn, will sharpen this notion even more. The provisional definition of amoralism in terms of lacking certain beliefs *and* acting in a certain way does not beg any theoretical questions. It is yet to be discovered whether the amoralist is able to believe or act otherwise, or whether his condition is the result of the psychological or physiological forces beyond his control. In the latter case, the question of the moral blame and responsibility of an amoralist will have to be considered. The possibility of the 'unwilling' amoralist would require to further modify Solomon's definition – the language of *rejection* of morality is no more meaningful when addressed to the mentally deranged maniac whose condition is caused (let us assume) by organic brain disorders, social conditions or hormonal disbalances. Rejection implies a voluntary conscious decision, and it might not always be the case that an amoralist is capable of such an act. Thus, it is preferable to talk about the *lack* of certain beliefs about morality and the lack of concern for morality, rather than the actual rejection of those beliefs. Furthermore, Solomon's insistence on *acting* "without regard for generally accepted moral principles" seems too strong as well. A passive amoralist (or the one in

[35] Alternatively, Jean Hampton (1989) suggests that the immoralist's behavior should rather be seen as a conscious *rebellion* against the moral injunctions, whose authority is categorical (i.e., an immoralist still understands that he is supposed to be governed by it.) For Hampton, an immoral act is primarily an act of *defiance*, and it is not necessarily done in view of some benefit that the wrongdoer may thus obtain.

confinement) might never have a chance to actually violate the moral norms in question; what is important is a strong *disposition* to do precisely that.

I have identified the existence of a *justifiable* moral system as a *minimal* requirement for the phenomenon of amoralism to have any philosophical weight, but nothing has been assumed about the *kind* of justification needed in order to 'ground' the moral system in the required manner, nor even whether there can be only one such system (or one justification strategy). It is conceivable that several distinct yet *compatible* moral systems can satisfy the justification requirement. I am assuming, however, that there can be no equally justified *incompatible* systems of morality.[36] The last assumption is needed to preserve the strong intuition that no one can be both moral and amoral at the same time and in the same respect.[37]

Solomon's definition of amoralism, which I cautiously accept with the above mentioned qualifications,[38] is wide enough to encompass a great variety of the amoral individuals, from the petty criminal who imagines himself being above the conventional morality to the extreme sociopath who is unable to comprehend even the most primitive forms of moral reasoning. Before we can start searching for some common thread in all these manifestations of amoralism, it will be necessary to further specify the 'sub-species' of this

---

[36] I.e., the systems whose basic non-derivative moral principles are inconsistent with each other.

[37] If a (justifiable) moral system A is incompatible with the (justifiable) system B, then there must be at least one pair of requirements, one from each system, which cannot both be satisfied at the same time. A person who accepts (and acts upon) the first requirement must necessarily reject (and fail to act upon) the second one. Thus, his act will be moral from the point of view of the system A, and amoral from the point of view of the system B.

[38] Perhaps another (obvious) qualification should be mentioned. The moral development of young children, it appears, does not reach any level of maturity till the age of four (see Piaget, J., 1965; Zahn-Waxler, 1992). It would be unhelpful, however, to classify children at the younger age as being amoral, given the strong negative connotations of this term. They are rather *pre-moral*.

general kind. But first our tentative definitions of amoralism and immoralism must be tested and defended against one of the experts in this area, Ronald Milo.

2.1 Ronald Milo's Typology of Immorality

Ronald Milo's seminal work on the concept of immorality and the typology of immoral behavior sets a high standard for any subsequent work on this subject.[39] Kurt Baier contends that "there is no comparable treatment of this topic in this century,"[40] and he is certainly right that even twenty-three years after its first publication there is no other systematic philosophical inquiry into the varieties of immoralism of the similar thoroughness and scale. If Milo's name has not been mentioned up till now it is because his book deserves a special treatment in a section of its own.

Since Milo's understanding of some of the basic concepts differs significantly from the one adopted in this work, I shall first present his division of immoralism and immoral behavior into the six different types, and state my main reasons for rejecting his specific taxonomy. I shall further explain how Milo's project is different from the one of the present work, and shall argue that many of his insights are nonetheless invaluable for any work that seeks to understand and explain the phenomenon of amoralism.

To begin with, Milo, unlike those philosophers discussed in the previous section, treats immoralism as an all-encompassing genus, where amoralism is just one of the subspecies. All amoralists are thus immoralists,[41] on Milo's classification, but not vice versa. He originally presents a twofold division of immorality: *wickedness* and *weakness*. These two cases of immoral behavior are also referred to as *unconscious* and *conscious*

[39] Milo, R. *Immorality*, 1984.

[40] Quoted in Milo, 1984, p. 1.

[41] With an exception of very young children, whom he (surprisingly) calls amoral but not immoral (pp. 59-60).

wrongdoings respectively. Wickedness (unconscious wrongdoing) is further divided into *perverse wickedness*, *moral negligence*, and *amoralism*; and the three subspecies of weakness (conscious wrongdoing) are *preferential wickedness*, *moral weakness*, and *moral indifference*.

*Preferential wickedness*, according to Milo, consists in an agent's doing what is morally wrong (while being conscious of its wrongfulness), because he (coldly) prefers the realization of some other end (usually defined in terms of narrow self-interest) to the morally praiseworthy alternative. In cases of *moral weakness* an agent realizes that what he does is morally wrong but due to his weak will, 'succumbs to temptation.' Presence of regret and feelings of guilt *after* the act are typical of this kind of immoral behavior as compared with preferential wickedness, where no such feelings usually occur. It also seems possible for a person to do something that he knows (or believes) to be morally wrong not as a result of a temptation or a conscious decision to pursue some other end incompatible with the morally praiseworthy behavior, but rather as a result of total indifference to the fact that the act is morally wrong. Milo refers to this form of immoral behavior as *moral indifference*.

A person is said to exhibit *perverse wickedness* if his immoral behavior is due to a mistaken belief that what he does is actually morally right. Following Aristotle, Milo recognizes the two basic ways in which a person may fail to believe that what he does is wrong: either he is ignorant of the correct moral principles (he mistakenly believes that the *perverse* action-guiding principles are the correct ones), or else, he is ignorant that his particular act falls under the type that is prohibited by the moral principle (e.g., he fails to appreciate that taking money from his father's wallet without asking falls under the category of stealing). When the immoral behavior results from the ignorance of the particular facts

(not general principles), and when a person is nonetheless blameworthy for his ignorance (he *should* have known better, even if he did not know better as a matter of fact), we have a case of *moral negligence*. Finally, the third type of unconscious wrongdoing is exemplified by a case of an agent who has neither correct nor perverse moral principles. When acting wrongly (as judged from the third person perspective), he neither believes that what he does is wrong nor believes that it is right. Milo refers to such pattern of behavior as *amoral behavior* or *amorality*.

How does Milo's detailed description of the six possible types of immoral behavior fits with our basic division of all non-moral behavior into immoralism and amoralism? Does our dichotomy coincide with Milo's root types of conscious and unconscious immorality (weakness and wickedness)? The answer is apparently 'no'. Even on the first reading of Milo's account it becomes obvious that his classification cuts across our division between amoralist and immoralist. An amoralist, on the definition adopted here, is a person who acts without regard for the generally accepted moral principles, e.g., a person in whom moral beliefs play no motivating or restraining role. We may say that his consistently immoral behavior shows (even if his verbal behavior contradicts it) that he does not in any real sense recognize the validity of moral principles, or their applicability to his own person. An amoralist lacks morality not in the sense that his morality is different from the morality of a society he lives in, but that whatever action-guiding principles he might have, they do not constitutive a system of *moral* principles.[42] Thus defined, our amoralist is like Milo's amoralist (unconscious wrongdoer) who has no normative moral beliefs at all. But he is also similar to Milo's conscious wrongdoer – a morally indifferent person, who knows the correct

[42] It will be the argument of Chapter II that there are constraints on the content and the form of action guiding principles in order for them to qualify as *moral* principles.

moral principles (has the correct moral beliefs), but fails to be motivated by them. Moreover, one can interpret Milo's examples of perverse wickedness as instances of amoralism in the wider sense, adopted here. If the perverse 'moral' principles adopted by a perversely wicked person cannot be properly called *moral* principles after all, we may legitimately say that such a person lacks morality in the required sense. Hence, amoralism as I shall use the term in this work will include at least three categories of immoralists from Milo's list.

On the other hand, the notion of amoralism adopted in the present work is also *narrower* than Milo's. Milo defines an amoral person proper as a person who does not possess the concept of moral goodness, and, as a consequence, when acting wrongly, does not believe that he acts wrongly but neither does he believe that he acts rightly. Given this definition, the behavior of very young children will fit this description (e.g., when a child sticks a pin into his baby brother), and Milo is ready to acknowledge that such a child is *amoral* (although not blameworthy).[43] I resist including young children in the category of amoral persons mainly because of the strong negative connotations that are part of the intension of amoralism. It is more natural, I suggest, to refer to young children as pre-moral.

Milo's description of preferential wickedness fits best with our understanding of immoralism, as long as the former is interpreted in such a way that 'knowledge' of (correct) moral principles that a preferentially wicked person allegedly has amounts to more than being simply aware of their existence. An immoralist must 'recognize' these principles in some stronger sense than simply knowing that others tend to observe them, and, in addition, consciously and systematically prefer to pursue some other end, incompatible with the end

[43] *Ibid*., p. 59.

of morality.[44] Finally, I regard moral weakness (weakness of will) and moral negligence as cases neither of immoralism nor amoralism proper. Even though immoral acts are committed in both cases, the strength of the 'temptation' and the ignorance of the particular facts may supply the extenuating circumstances for the agent. The distinguishing marks of moral weakness is the presence of guilt or remorse after the act, *and* the fact that such 'moral lapses' are rare and far between. The distinguishing mark of immoral conduct due to moral negligence is that it can usually be easily remedied by *moral instruction*, i.e., supplying the agent with the true factual beliefs relevant to his act.

The main reason why my definitions of amoralism and immoralism are preferable in this case to Milo's more detailed classification is that my twofold distinction allows greater flexibility in discussing the particular instances of various types of immoral and amoral behavior without making the controversial and costly assumptions about certain meta-ethical theories. Milo acknowledges that the six types of immorality that he described are *factually* possible only if *externalism* is true about moral motivation, *cognitivism* is true about the nature of moral beliefs and *descriptivism* is true about the material and formal constraints on moral beliefs (as opposed to non-moral ones).[45] The broader (and less restrictive) definition of amoralism that is adopted here, however, requires, strictly speaking, only the truth of the last condition – descriptivism about the nature of morality (which will be argued for in the

[44] I am aware of the problem of specifying the sense in which an immoralist 'recognizes' the moral principles (or the authority of these principles) but yet remains unmoved. However, the problem of immoralism is not my main subject matter. If it turns out that strong motivational internalism is true, and no genuine immoralists are possible, nothing will be lost for this project.

[45] *Ibid*., pp. 14-17. E.g., if strong motivational internalism is true, then moral indifference and preferential wickedness are impossible. Likewise, moral indifference will be impossible if descriptivism is wrong, and one holds instead that "*any* principle can be accepted as a moral principle if one is prepared to accept it as overriding." (p.15). Furthermore, moral *non-cognitvism* will make moral indifference, moral weakness and preferential wickedness paradoxical.

next chapter). But the actual existence of amoralists is compatible with the non-cognitive analysis of moral judgments as well as with most versions of motivational internalism.

Yet Milo's typology seems warranted in view of the specific and limited goals of his book, of which he speaks in the beginning:

> My chief concern in this book is with the typology of immorality. […] No attempt will be made to describe how human beings actually behave or to classify their behavior in such a way as to gain a better psychological understanding of it. Rather, the purpose here is to come to a better understanding of the concept of immorality. Thus, I shall be concerned with the logical possibility of various types of immoral behavior, rather than with their instantiation or incidence.[46]

On the contrary, the conceptual issues will be only of secondary importance in the present investigation, and a special effort will be made to point to the real-life instantiations of amoralism. Since it is unlikely that all six types of Milo's immoralists actually exist or ever existed in their pure form, a more liberal and necessarily less specific description of amoralism is adopted here. But even with the most inclusive understanding of amoralism and immoralism, it is important to recognize that few (if any) actual cases will *perfectly* exemplify these abstract philosophical categories. In the vast majority of instances, we will deal with partial instantiation only.

Whereas Milo's interest lies primarily in the analysis of the *concept* of immorality, I consider the development of an *explanatory* theory of the actual instances of amoralism to be my main goal.[47] Milo's discussion of the *logical possibility* of the various types of amoralism will yet prove invaluable in this work as well. In what follows, I shall rely on a number of

---

[46] *Ibid*., p. 26.

[47] Milo does devote a small section to the question of the roots of immorality, but acknowledges that this is largely a side issue for his project.

arguments and insights from Milo's work, and employ parts of his rich terminological apparatus for my own purposes.

3. The Degrees of Amoralism

Once we define an amoralist as an individual who fails to recognize the validity and overriding importance of moral requirements,[48] and, furthermore, fails to consider the moral reasons in his everyday actions (or the one who is disposed to make choices regardless of the considerations of morality), it becomes evident that being amoral is a *matter of degree*. As Jeffrie Murphy observes "we all have our psychopathic tendencies, and so all the actual cases may be neither black nor white but various shades of gray."[49] The following classification of amoralists into two logically distinct groups or classes is just *a* way to delineate this complex social phenomenon, but it is by no means the only possible one. It is also unlikely to be an exhaustive grouping. It will be acceptable, however, as long as it allows us to place definitively the most notorious cases of amoralism in one of these two categories, even if a number of the problematic borderline cases will still remain.

In order to set the visible limits to the degree of amoralism right from the start, we should begin with the most extreme cases. As in so many other instances, here, too, Aristotle provides a starting point.

In the *Nicomachean Ethics* Aristotle distinguished between the virtuous, continent, incontinent, vicious and the brutish characters.[50] From the brief description given of the last

[48] The non-cognitive version of the same definition will stress the lack of a *pro-attitude* toward certain types of action, which we would consider the morally right actions.

[49] Murphy, J. 1972, p. 296.

[50] Aristotle, *Nicomachean Ethics*, Book VII, 1145a15 – 1145b5.

character, the brutish type, we may gather that Aristotle had in mind a kind of person who closely resembles the extreme amoralist that interests us in the present work. He introduces the brutish character in the following passage from Book VII:

> Now, since it is rarely that a godlike man is found, so too the brutish type is rarely found among men; it is found chiefly among barbarians, but some brutish qualities are also produced by disease or deformity; and we also call by this evil name those men who go beyond all ordinary standards by reason of vice.[51]

There are several features that Aristotle takes to be definitive of the 'brutish type'. He states that "a brute is a different kind of state from vice"[52] and further specifies that "those men go beyond all ordinary standards by reason of vice."[53] Later in the same book he remarks that "wickedness which is on the human level is called wickedness simply, while that which is not is called wickedness not simply but with the qualification 'brutish' or 'morbid'"[54], and mentions as examples of such nonhuman wickedness the graphic stories of a female who rips open pregnant women and devours the infants, the savage tribes along the Black sea that give their own children to feast upon, and the like.[55]

A more formal definition is given when Aristotle classifies the brutish characters as "people who are thoughtless and live by their senses alone."[56] We may understand this as referring to people who are unwilling or unable to restraint their natural or, rather, 'unnatural'[57] sensual impulses by the motivating power of practical reasoning - a kind of

---

[51] *Ibid*., 1145a15-b05.
[52] *Ibid*., 1145a26.
[53] *Ibid*., 1145a33.
[54] *Ibid*., 1149a16-17.
[55] *Ibid*., 1148b23ff.
[56] *Ibid*., 1149a9-10.

[57] Aristotle argues that certain patterns of conduct and certain feelings that are typical of the brutish characters cannot be "pleasant by nature" but their attractiveness to the agent must result either from a disease, acquired through perverted habits, or else originates "by reason of originally bad natures". (*Ibid*., 1148b16-19)

reasoning that corresponds to the intentional actions of a virtuous agent. The immediate gratification or pleasure expected from committing 'immoral' acts is the only motivating reason operative in these individuals. Such self-indulgence, however, must go beyond the one that is characteristic of an incontinent man, and a brutish person may be called incontinent with respect to the excessive state he is in by analogy only.[58] It is an important point for Aristotle that there can be no 'mean' to the desires and feelings of a brutish person: their very presence is the result of some perversion of human nature, and that is what puts the brutish type beyond the realm of simple vice, or occasional incontinence.

Aristotle's description of a 'brute' closely resembles the description of an aggressive psychopath or sociopath[59] as given by modern psychologists and criminologists.[60] One of the first modern researchers on the phenomenon of sociopaths (although not under this name) was English physician and ethnologist James C. Prichard, who published an important "Treatise on Insanity" in 1835. He referred to the antisocial sociopath as a 'morally insane' person and described this condition in the following way:

> There is a form of mental derangement in which the intellectual faculties appear to have sustained little or no injury, while the disorder is manifested principally in the state of feelings, temper, or habits. The moral and active principles of the mind are strangely perverted and depraved; the power of self-government is lost or greatly impaired; and the individual is found to be incapable, not of talking or

---

[58] *Ibid*., 1149a2.

[59] Many scholars who work in this area use the terms 'sociopath', 'psychopath', and 'person with antisocial personality disorder' interchangeably. Indeed, the latter phrase is usually preferred as the more comprehensive one, and the former two terms are somewhat fading away in the recent literature. (Although see Lykken (1995) and Stout (2005) for insisting on the important difference between these types). I shall generally follow this convention except in cases where greater precision is needed.

[60] For the argument that modern 'brutish characters', i.e., serial killers and maniacs, are persons falling under the category of psychopaths see Giannangelo, Stephen J. *The Psychopathology of Serial Murder: A Theory of Violence*. London: Praeger series, 1996.

> reasoning upon any subject proposed to him, but of conducting himself with decency and propriety in the business of life.[61]

Aristotle's above description of the brutish characters as "thoughtless people [who] live by their senses alone" may plausibly be interpreted as referring to the phenomenon of the "loss of self-government" in Prichard's sense.[62] Even though it is not clear how much actual empirical observation of the 'brutes' was done by Aristotle himself (if any) and how much of his data was taken from the anecdotal evidence of the popular ghastly stories about the 'barbarians', widely spread in the Greek world at that time,[63] still both Aristotle and Prichard seem to be pointing in the direction of a certain kind of individuals with quite distinctive behavioral and cognitive traits who would later be referred to as psychopaths or sociopaths.

Since the first attempts at classification and description made by Prichard in the beginning of the nineteenth century, scholars have added a number of additional features to the list of the common personality and behavioral traits characteristic of the psychopaths. The poor behavioral control, the inability to resist one's immediate desires and impulses, the absence of the usual internal constraints (e.g., shame, fear of punishment, guilt), lack of empathy and unusual cruelty to other people and animals are just some of the defining features of sociopaths appearing in the most recent *Diagnostic Manual,* as well as in the more informal descriptions of those suffering from antisocial personality disorder made by individual scholars.[64] Thus Wolman argues that "the 'pleasure principle,' that is, the principle

---

[61] Prichard, J., 1973, p 15.

[62] Although Prichard's notion of 'self-government' might encompass more than simply the power of practical reasoning (as for Aristotle), but also the faculty of conscience as a distinct 'governing' organ (a characteristic trait of the old 'faculty psychology').

[63] For more reliable reports of the 'savage tribes' see Colin Turnbull's "The Mountain People" (1987)

[64] *Diagnostic and Statistical Manual of Mental Disorders-IV*, (DSM-IV), 1994. Also: Lykken (1995); Hare, (1996), Wolman (1999), et al.

of immediate gratification of needs, is the main motive in the life of sociopaths."[65] And David Lykken, in addition, notes that "psychopath is characterized by a lack of the restraining effect of conscience and of emphatic concern for other people."[66] The absence or unusual weakness of conscience in psychopaths, the faculty traditionally associated with moral guidance and the ability to discern right from wrong, is also emphasized by Robert Hare, one of the foremost authorities in this area:

> For psychopaths, the social experiences that normally build conscience never take hold. Such people don't have an inner voice to guide them; they *know* the rules but follow only those they choose to follow, no matter what the repercussions for others. […] Without the shackles of a nagging conscience, they feel free to satisfy their needs and wants […] Any antisocial act, from petty theft to bloody murder, becomes possible.[67]

Even these brief descriptive remarks make it clear that a psychopath falls under the category of amoralism as defined in the previous section.[68] Individuals suffering from the antisocial personality disorder cannot be said to accept or recognize the moral norms and principles prevalent in their society in any plausible sense of these terms. Accepting a rule that purports to be action-guiding (as moral principles surely do) implies at least minimal behavioral manifestation of such acceptance. Since the behavior of the most psychopaths (even the non-aggressive ones) exhibits no signs of being in any way influenced or altered by these norms (except accidentally, e.g., in cases where the requirements of morality happen to coincide with some immediate impulse of a psychopath), we may conclude that they *lack*

---

[65] Wolman, B., 1987, p. 44.

[66] Lykken, D., 1995, p. 115.

[67] Hare, R., 1993, p. 75-76.

[68] The label of *amoralist* is attached to a psychopath (or sociopath) by Norman Williams (*Introduction to Moral Education*. Baltimore: Penguin Books, 1967, p. 272), Derek Wright (*The Psychology of Moral Behavior*. Baltimore: Penguin Books, 1971, pp. 208ff), and David Brink ("External Moral Motivation" *Southern Journal of Philosophy*, 1986, p. 29). For the contrary opinion see Milo, 1984, pp. 60-62.

morality in the strongest sense. The persistent disposition to act contrary to the moral constraints and without any regard for other human beings is a consequence of this lack.

I suggested earlier that it is unhelpful to limit the class of amoralists to psychopaths only (as Williams prefers). A psychopath might figure as a paradigm case of amoralism, but he need not to be the only type. For convenience sake, I shall refer to a person who closely resembles the sociopathic personality as described in the recent literature, or a 'brutish character' and a 'morally insane' person as described by Aristotle and Prichard respectively, by the name of *extreme* or *thoroughgoing amoralist*.[69] Identifying these cases will both set the limit to the degree of amoralism and will allow me to define the less severe cases against the visible background of these extreme instances.

Besides the thoroughgoing amoralist, there might be a person who, without *fully* abandoning the moral point of view, limits severely the *scope* of such a disinterested concern for others as inherently required by this point of view.[70] Such selective application of moral considerations may be either persistent or temporary. Kai Nielsen, for instance, describes a 'classist amoralist' as a person who "only extends his disinterested caring to his own peers [members of the elite class], but who treats all other people manipulatively, deploying morality as moral ideology to keep those people in line."[71]

But such arbitrary limiting of the scope of morality may also temporarily occur under the extreme conditions of hostilities, famine, social unrest and the like. Jonathan Glover in

---

[69] The category of the extreme amoralists need not be fully coextensive with the medical category of psychopaths. Some autistic individuals, for instance, may very well lack morality in the relevant sense, and can thus be described in similar terms (See Grant, 2005).

[70] I shall argue in Chapter Two that a disinterested concern for others is at least partially constitutive of the moral point of view proper.

[71] Nielsen, Kai., 1989, p. 295.

"Humanity: A Moral History of the Twentieth Century" describes situations such as wars, local military conflicts and social revolutions where normal behavioral restraints break loose.[72] Individuals, who were exemplary members of society with high moral standards, may turn into ruthless murderers in the time of war. It is implausible to suggest that these people temporarily 'forgot' the requirements of morality, nor is it likely that a totally new set of moral principles (or 'immoral' ones) was acquired for the period of hostilities. But just as with Nielsen's 'classist amoralist', in this case too, the scope of people who fall under the domain of moral obligation is arbitrarily limited, even if only for a certain period of time: the appropriate moral requirements and obligations, for instance, may apply to people on this side of the 'front line', but not to people on the other side. The most conspicuous cases of this kind of amoral attitude have eerie historical reality to them. The following excerpt from the speech made by Heinrich Himmler to SS Group leaders in October of 1943 illustrates this especially well:

> It is absolutely wrong to project our own harmless soul with its deep feelings, our kindheartedness, our idealism, upon alien peoples. […] One principle must be absolute for the SS man: we must be honest, decent, loyal and friendly to members of our blood and to no one else. […] Whether the other races live in comfort or perish of hunger interests me only in so far as we need them as slaves for our culture; apart from that it does not interest me. We shall never be rough or heartless where it is not necessary; that is clear. We, Germans, who are the only people in the world who have a decent attitude to animals, will also adopt a decent attitude to these human animals.[73]

[72] Glover, J., 1999.

[73] Quoted in Bennett, J. (1974), pp. 127-8.

The *selective amoralist,* as we may now call him, is a member of the same species, even though, of course, he is not wholly devoid of morality in the way a psychopath is[74] (e.g., Himmler recognizes the demands of decency, honesty and loyalty to the members of one's race). His amoralism, rather, is manifested either by persistent unwillingness to recognize a certain group of people as having all the rights of the other moral agents or by limiting the extent of his moral duties to the time and place that are chosen arbitrarily.[75]

What these two varieties of amoralism have in common is that these individuals have abandoned (or never entered) the common "world of morality", in B. Williams' phrase. Since such an amoralist is clearly not a merely fictional character, an explanation and further elucidation of this phenomenon is needed. This dissertation is an attempt to provide such an explanation. But in a philosophical work we are not looking for a straightforward *causal* (anatomical) explanation of why some people ignore morality. The philosophically interesting question is the one that asks for the rational justification of the moral point of view in its most general features, and then considers whether the peculiar deficiencies of the amoralist who is unable (or unwilling) to appreciate such a view, are due to some cognitive mistake in reasoning, or the lack of some emotive gear which constitutes a necessary condition for being in the realm of morality. The successful discovery of the explanatory

[74] I have indicated above that, following convention, the terms 'psychopath' and 'sociopath' are used interchangeably in this work. However, some writers suggest to draw the distinction between the two classes along the following lines: whereas the term 'psychopathy' is reserved for the most severe cases of amoralism, sociopathy, on the other hand, "refers to criminal attitudes and behavior viewed as normal in certain groups, such as street gangs; sociopaths have a sense of right and wrong that is based on the values of their criminal group." (Bower, Bruce. "The Predator's Gaze." *Science News*, vol. 170, 2006, p. 379). I can accommodate this terminological discrepancy. My distinction between *extreme* and *selective* amoralists would then (roughly) correspond to this proposed distinction between psychopaths and sociopaths.

[75] 'Arbitrarily' does not mean randomly. A person may choose to respect moral duties only in cases where moral obligation does not go against his perceived self-interest, or only when it is not true that committing an immoral act can bring great benefits (cf. Hume's 'Sensible Knave' and Hobbes' '*Foole'*). Clearly, the choice of when and where should one be moral is not random in those cases. But it is an arbitrary selection from the moral point of view, which inherently demands a certain level of abstraction from one's particular circumstances and private interests (as I shall argue in Chapter Two).

reason (or reasons) of amoralism will provide a better understanding of what morality needs in order to get off the ground, and what features are normally critical to the development of a moral agency.

3.1 Kant's Degrees of Wickedness

The discussion of this section would be incomplete without mentioning at least one alternative classification of amoralists with respect to severity of their amoral condition. The topic of gradation of the degrees of amoralism has received an unduly scant attention from the philosophers, yet one conception, coming from Kant's writings, is worth mentioning here. In what follows I shall summarize Kant's view on this matter and offer reasons for rejecting his classification as intuitively unsatisfactory and, in general, unhelpful.

Kant distinguishes three different degrees of the "propensity for evil" in his "Religion Within the Boundaries of Mere Reason", and all three are related to the failure to adopt good maxims.

> *First*, there is the general weakness of the human heart in complying with the adopted maxims, or the *frailty* of human nature; *second*, the propensity to adulterate moral incentives with immoral ones (even when it is done with good intention, and under maxims of the good), i.e., *impurity*; third, the propensity to adopt evil maxims, i.e., the *depravity* of human nature, or of the human heart.[76]

The last and the highest degree of moral failure, depravity or wickedness, is then further explained in the following passage:

> The *depravity* or, if one prefers, the *corruption* of the human heart is the propensity of the power of choice to maxims that subordinate the incentives of the moral law to others (not moral ones). It can also be called the perversity of the human heart, for it reverses the ethical order [of priority] as regards the incentives of a *free* power of choice; and although with this reversal there can

[76] Kant, I. (1793), 1998, p. 53 [6:29]

> still be legally good actions, yet the mind's attitude is thereby corrupted at its root, and hence the human being is designated as evil.[77]

How does Kant's gradation of morally inadequate persons relate to our discussion of amoralism and related notions? First, it is natural to identify Kant's *frail* human being with the traditional notion of an incontinent or a weak-willed person. In the classic Aristotelian account of incontinence,[78] the incontinent person accepts at least the major premise of the practical syllogism, which expresses the general moral rule. His universalized maxims, in Kant's terminology, are logically consistent and thus morally good. But the failure to act on the correct moral rules is then explained by the inability to identify the particular situation as falling under the relevant general category of the major premise, and this temporary 'ignorance' or 'blindness' is in turn attributed to the overpowering force of the competing practical reasoning, where desires and other non-cognitive elements play a leading role.

Secondly, a person with *impure* maxims (or impure heart) would acquiesce to the demands of moral law and would act in accordance with such demands, but only with the help of some non-moral incentives (e.g., self-interest). A prudent tradesman who deals honestly with his customers *both* because it is his duty and because it is good for business would illustrate such impurity of motives. Such a person, according to Kant, is not yet wholly depraved to the extent that at least some of his actions proceed from the 'good will', i.e., from the motive of doing the right thing for its own sake.

Finally, there is the notion of a perverse or evil person, and here Kant's rhetoric becomes much more powerful. It is tempting to interpret wickedness ('perversity of heart') as a notion coextensive with amoralism, and thus falling under the scope of the present study. Yet Kant's characterization of the 'third degree' of the capacity for evil raises a number of

---

[77] *Ibid*., p. 54 [6:30]
[78] *Nicomachean Ethics*, Book VII, Chapter 1-3.

questions. In the second quote above Kant defines wickedness as the "the propensity of the power of choice to maxims that subordinate the incentives of the moral law to others (not moral ones)." In other words, a perverse (wicked) person tends to adopt non-moral maxims as his subjective principle of volition, and acts on the basis of such maxims. But his acting on the 'wrong' incentives must be more thoroughgoing and systematic in order to distinguish him from a person with impure heart. However, Kant states that the external actions of a person with a "corrupted heart" may quite possibly be "legally good", i.e., be in accordance with one's moral duty. The famous "philanthropist" from the "Groundwork" who helps others out of sympathetic concern for their needs is clearly acting on a non-moral maxim, even though his conduct conforms to duty. Still, as it follows from Kant's analysis, such a "friend to humanity" is an example of the greatest degree of corruption insofar as he develops a stable disposition to act upon these non-moral maxims, and ignore the properly moral ones. To be sure, the worst moral monster one can imagine will fall in the category of the perverse people as well, according to Kant. But the key here is that one does not *have to* act contrary to moral requirements, for Kant, in order to deserve the label of an evil person: "We call a human being evil, however, not because he performs actions that are evil, but because they are so constituted that they allow the inference of evil maxims in him."[79]

But such unprecedented widening of the category of moral wickedness is hard to understand let alone to justify. Richard Bernstein, discussing this passage, is quite right when he says that

> To judge such a person [i.e., a sympathetic philanthropist] as the paradigm of *wickedness*, to put him in the same category as a mass murderer – at least, in respect to the degree of evil exhibited – is more than an awkward consequence

[79] *Ibid*., p. 46 [6:20]

of Kant's rigorism; it is morally perverse.[80]

The unwarranted inclusiveness of Kant's category of moral perversity greatly diminishes its significance for moral theory. As I argued above, the notion of amoralism, if it is to remain a meaningful notion at all, must be defined in a way that sets strict limits to its application. Moreover, the conditions of the concept's application must be at least partially empirically verifiable. Stipulating actual systematic immoral behavior (acting contrary to moral duty) as part of what it means to be an amoral person was an essential part of the intensional characterization of amoralism. As we have seen, Kant's notion of *perversity* or *wickedness* is much wider than amoralism thus understood.

I conclude that Kant's classification of amoral or evil persons with respect to severity of their condition is not a real competitor to the one proposed above. For the purposes of the present investigation, the distinction between the extreme and selective amoralists seems to be, at this stage, the most optimal one, and its usefulness will be further demonstrated in the course of the argument.

## 4. The Outline of the Project

The present study is an interdisciplinary work that employs the resources of moral philosophy, clinical psychiatry, literary criticism and moral psychology, and seeks to present a comprehensive picture of the phenomenon of amoralism in its various manifestations with the goal of understanding its underlying causes. A successful explanatory theory of amoralism will yield further insights on the nature moral agency, the conditions of moral performance and, potentially, provide a basis for the full-fledged theory of moral motivation.

---

[80] Bernstein, R. 2002, p. 29.

Since the amoralist was defined as someone who ignores or willfully abandons the moral point of view, we must first try to list and justify the constitutive features of morality in general without assuming the truth of any particular system in normative ethics. Some formal requirements (e.g. universalizability), as well as constraints on the content (e.g., the moral principle must deal with the interests of others) will be presented as the indispensable features of *any* system of morality. I shall argue for the truth of *descriptivism* against *neutralism*, and discuss the four constraints on the application of the term 'moral' to an evaluative judgment, and the term 'morality' to a normative action-guiding system. The resulting picture will provide a necessary contrast *against* which we can further define and elucidate a non-moral standpoint of the amoralist. One of the conclusions of this part will be that rational individual egoism and a number of similar views that arbitrarily select either one individual or a group of people as the only ones whose interests ought to be considered are not simply alternative, although peculiar, moralities but rather non-moral policies altogether. Thus anyone adopting such a policy (a non-moral action-guide) will constitute a clear example of an amoralist in the required sense.

In the course of building my argument in the third chapter I shall consider in much greater details the various empirical manifestations of amoralism as it was introduced and defined in the previous section. I shall draw examples from the anthropological data, medical records, literary characters and historical cases. An attempt will be made to identify some common feature or features that all amoralists share. This will prove crucial for the subsequent parts of the dissertation which will inquire into the common ground of all cases of moral indifference.

The two general explanatory possibilities will be further considered. If morality is essentially a mode of human behavior which can be *rationally* justified all the way down, it must necessarily be the case that anyone failing to assume a point of view of morality acts contrary to the requirements of reason as such. An amoralist, on this (roughly 'Kantian') reading, commits a cognitive mistake and can be properly accused of *irrationality* (even if this will not move him to a slightest degree to become moral). If, on the other hand, the basic underlying features of morality have emotive rather than purely intellectual roots, a person who lacks or suppresses the required sentiments will be forever unable to appreciate morality in the full sense of the word. The accusation of *insensitivity* might be then in order, yet, this will hardly point to a flaw in a person's cognitive abilities.

After carefully considering and analyzing these possibilities (in Chapters IV and V), I shall defend the following conclusions. While recognizing the importance of emotions (e.g., empathy) for moral performance, it will be argued that having a sympathetic concern for another (the unimpaired 'emotive apparatus') is not the *only* way into morality. The examples of autistic individuals who lack empathy but yet are able to acquire full moral competence will be discussed in this connection. The attention thus must be shifted to the ways the amoralists judge and think rather than feel. The close analysis of Ted Bundy's reasoning process, a notorious serial killer, will prove invaluable in this context. An attempt will be made to show that the only way to preserve the sense in which the behavior and reasoning of amoralists can be meaningful judged as abnormal (and thus rationally defective) is by extending the domain of rational assessment of action from the purely instrumental application to the final *ends* of action and the substantive values that guide one's behavior. Robert Audi's general analysis of practical rationality, its scope and structure, will be

employed in disclosing the dynamics of morally relevant decisions and determining the specific cognitive failure of amoralists. It will be argued that mistaken value judgments, including the judgments establishing the relevant priority of values, are at the core of the amoral condition.

In the concluding chapter (Chapter VI) I shall probe the limits of the explanatory efforts, and provide theoretical reasons for the claim that amoralism must, in the final analysis, remain unexplained. The final conclusion, thus, will be largely negative. The *ultimate* reason why some people choose to ignore the moral constraints and systematically act without references to the demands of morality cannot be fully specified. The explanatory chain must come to a forced stop when we encounter the sphere of human freedom of choice. This fundamental limitation to our knowledge has both a positive and a negative aspect. On the negative side, it limits our ability to fully understand the underlying deep causes of amoralism, and, hence, successfully predict, prevent and treat the particular cases of amoral conduct. On the positive side, however, the inscrutable ground of free human choices that form the amoral character and lead to amoral actions supplies the necessary conceptual framework for holding the amoralists fully responsible for who they are and what they do.

# CHAPTER II

## THE MORAL POINT OF VIEW

> I never approve, or disapprove, of anything now. It is an absurd attitude to take toward life. We are not sent into the world to air our moral prejudices. I never interfere with what charming people do. If a personality fascinates me, whatever mode of expression that personality selects is absolutely delightful to me.
>
> Oscar Wilde, *The Picture of Dorian Gray*
> (Lord Henry Wotton)

The amoralists were defined in the previous chapter (partly) in terms of *lacking morality*. It is now prerequisite for the success of the explanatory project that we first clearly state what constitutes the moral point of view proper, i.e., the point of view that an amoralist is said to reject, ignore or is simply unable to adopt. A discussion of the normative repercussions of rejection of morality and why having morality is not optional will follow. Finally, we must explicate the various senses in which someone might *lack* morality or fail to adopt a moral point of view.

As to the first point, it will be argued that any adequate definition of morality must necessarily include reference to the *form* and *content* of the moral principles and moral beliefs. As to the second point, I shall adopt a minimalist view of non-reductive autonomy of moral evaluative concepts. As to the last point, an argument will be advanced to show that the notion of 'lacking morality' must be understood primarily in terms of manifested indifference toward the non-arbitrary ends of morality, which ultimately can be traced to a mistaken value judgment.

1. The Problem of Demarcation: Neutralism vs. Descriptivism

The clear conception of the moral sphere will allow us, first of all, to demarcate the moral issues, rules, principles and prohibitions from the various common non-moral behavioral guides. Imagining oneself in a position of an anthropologist who discovers a previously unknown exotic culture with a complex system of social relations and a well-developed (but very unusual) code of the socially acceptable behavior patterns, we can readily appreciate the difficulty of sorting out the many rules that the members of that society actually observe or the ones they claim to be important into the moral and non-moral categories. What should we take as the criterion of a *moral* rule? We may call it the *problem of demarcation*[81] between the moral and non-moral spheres.[82] 'The aberration of closeness' generally hides this difficulty when we consider the more familiar social arrangements (such as our own society), but the problem is immediately felt as soon as we attempt to give a precise formulation of the distinctive features of the moral rules for conduct as opposed to the non-moral ones.[83]

On the most general level, one may arrive at the two different kinds of *substantive* theories of what constitutes the moral point of view proper. To begin with, one may emphasize the close connection between the normative rules that prompt or prohibit certain types of behavior and the actions of agents who claim to accept those rules. On this view, the

[81] The phrase 'demarcation problem' used in the present context first appears in Catherine Wilson's "Moral Animals" (2004), Chapter 1.

[82] This problem of demarcation equally applies to moral *rules* and moral *issues* as well as to moral evaluative *judgments*. We may ask what makes the evaluative statement "This is a bad weapon" an instance of a *non-moral* judgment, as opposed to "He is a bad person", which (presumably) is a moral judgment. We will still need to have some criterion to be able to make these distinctions. For brevity sake, I shall talk mostly about moral rules and principles but much of the same reasoning will apply to judgments as well.

[83] We can mention the rules of prudence (e.g., "Always stay away from stray dogs"), the rules of etiquette (e.g., "Always hold your fork with the left hand") and the game rules (e.g., "Move your pawn in such and such a way") as examples of the (prima facie) non-moral rules.

problem of specifying the *moral* rules (as opposed to the various non-moral guides) is a matter of observing which ones the members of a given society *actually* take as overriding, i.e., as taking unquestionable precedence in cases of conflict with other (non-moral) requirements. Morality, then, becomes coextensive with any action-guiding code of conduct adopted by an individual or a group of people. And the problem of recognition of moral principles in a society is thus reduced to the problem of identifying those rules (whatever the content) that are actually observed and honored by the majority of the members.[84]

I shall follow Kai Nielsen in referring to the position just introduced by the name of *neutralism*.[85] Neutralism in its pure form does not build any intrinsic constraints on the form or content of the moral principles, i.e., it remains neutral with respect to what those principles can be about or how they should be formulated. It is rather primarily concerned with the functional, action-guiding character of those rules, and makes the fact of actual observance of the rules by the agents into the *constitutive* part of morality.[86] Jonathan Bennett seems to advocate this position when he writes the following:

> There could be dispute as to whether the springs of someone's actions constitute *morality*. I think, though, that we must admit that someone who acts in ways

[84] As I have suggested above, this *functional* distinguishing mark of moral principles (namely, "whatever rules are actually observed" or "whatever is valued") can be arrived at (at least there have been *attempts* to that effect) both by the analysis of certain concepts (as when Hobbes, for instance, analyzed the 'good' as 'the object of desire'), and by the empirical approach. One can, for instance, observe that the core moral rules, which he takes as a given, exercise significant influence on behavior of the members of society, and infer from this that this is precisely what distinguishes the moral rules from the non-moral ones.

[85] Nielsen, Kai., 1989.

[86] It is possible to give both the *objective* and *subjective* formulations of neutralism. In the former case, it would be defined (as I did) in terms of empirically observed behavior of the agents, while in the latter case, the emphasis is given to the inner attitude of approval or disapproval that an agent acquires when presented with a principle of conduct or an occasion for a certain type of behavior (e.g., 'X is wrong' is defined as 'X rouses indignation in me'). As long as there are no references to content, form, goals or grounds of these attitudes, and 'whatever principle is (strongly) approved' is used as the criterion by which we distinguish between the moral and the non-moral, we still have a version of neutralism. Since it is natural to expect that the attitudes of approval or disapproval will be (generally) manifested in one's behavior, I shall mainly refer to conduct here.

> which conflict grossly with our morality may nevertheless have a morality of his own – *a set of principles of action which he sincerely assents to*,[87] so that for him the problem of acting well or rightly or in obedience to conscience is the problem of conforming to *those* principles.[88]

We can further illustrate neutralism as a philosophical position by reference to a relatively recent and influential tradition. J. P. Sartre in his earlier existentialist writings advocates an ethical theory that ties the moral status of an action to a free, unimpeded choice by an individual.[89] There is no natural or supernatural fact, according to Sartre, that should constraint the use of the concept 'moral goodness'. It would be acting in 'bad faith' to justify one's decision, say, to go to war, by reference to some special feature of an action, its foreseeable consequences, an abstract moral principle or concrete advise by another. An action, whatever it is, becomes morally right (and morally relevant) precisely because it is freely chosen and performed by the agent in the absence of any constraints of that kind. Whatever external features of an act are there, it is our free *endorsement* of these features that explains the action, rather than the features themselves. On this view, however, morality ceases to have any definite content, and there can be no meaningful distinction between correct and mistaken use of moral predicates, as long as they are sincerely used. Any type of action or a state of affairs with any set of objective features can become morally relevant through the performative process of endorsement.

One of the consequences of adopting neutralism as a kind of substantive theory of the moral point of view is that we can now legitimately speak of the 'Nazi morality', 'Slave-

---

[87] My emphasis.

[88] Bennett, Jonathan, 1974, p. 125.

[89] Sartre, J. P., 1957, pp. 9-40.

owner's morality', the 'Ik morality' and the 'Dobu morality'[90] (among others), without being accused of committing a category mistake. However perverse or unjustified were some of the racist principles practiced, for instance, by the Nazis, they were still the kind of principles that the Nazis took seriously and were implementing consistently in their conduct – i.e., the kind of principles that they, using Bennett's phrase, "sincerely assented to." As such, the set of *those* principles would constitute a peculiar *morality* of the Nazis, even if an extremely *bad* morality from our own perspective.[91]

It is now apparent that neutralism, if accepted, would make the notion of amoralism all but empty. Indeed, it would deny the possibility of amoralism. The aggressive psychopath whose anti-social behavior can perhaps be subsumed under the principle "Always follow your immediate impulse" or the like, would not be devoid of morality altogether, according to neutralism, but would rather have an unusual moral code of his own. As Richard Garner remarks in criticizing this view,

> Someone might say that the mere fact that we choose one course of action over another shows that we have moral principles. When "amoralists" behave as if they subscribe to moral principles, we can conclude that they do subscribe to moral principles, perhaps without realizing it.[92]

The actual behavior of extreme or selective amoralists can always be described as being in accordance with the rules and principles that they take as overriding. The principle of ethical egoism or any other principle that arbitrarily selects the group of people whose

---

[90] The anti-social and uncooperative patterns of behavior of the Ik culture in Uganda are documented by Colin M. Turnbull in *The Mountain People* (1987). The peculiar behavioral code of the Dobu culture of New Guinea was described by Ruth Benedict in her *Patterns of Culture* (1936).

[91] A short excerpt from Himmler's speech given in 1943 in front of SS audience may illustrate this: "We had the moral right vis-à-vis *our* people to annihilate *this* people which wanted to annihilate us. But we had no right to take a single fur, a single watch, a single mark, a single cigarette, or anything whatsoever…" (quoted in Vetlesen, 1994, p. 110). What is peculiar about this SS *Weltanschauung* is that the theft of a cigarette is morally wrong, but the collective annihilation of millions is part of one's moral duty.

[92] Garner, R. 1994, p. 283.

welfare is to count for more than the welfare of the outsiders is the kind of principle that can be held "sincerely" and manifested in one's everyday conduct. Those accepting these principles would satisfy the behavioral and dispositional requirements set by neutralism, and can in fact be considered *moralists* (as long as they do not deviate from the kind of behavior sanctioned by the adopted principles) by the people of similar moral convictions. Likewise, it follows that a dutiful SS officer who would help a Jew (e.g., out of compassion) would be committing an *immoral* act.

The last observation, however paradoxical, would not by itself discredit neutralism. Jonathan Bennett, for instance, considering the case of Huckleberry Finn and his predicament of what to do with a runaway slave Jim, is ready to admit that Huck's decision not to hand in Jim to the authorities was immoral from the point of view of *that* society, but was a morally praiseworthy action from the contemporary point of view[93] (which, I take it, he does not assume to be absolute or incorrigible). Once the moral point of view is relativized to the social circumstances and the specifics of the time period of an action, the conclusions like the one about the SS officer does not seem implausible after all.[94]

However, the initial plausibility of neutralism depends to a large extent on the plausibility of the assumption that all alternative attempts at specification of the moral point of view that take into account the form and the content of moral principles, and that seek to establish a core set of non-relative moral requirements and objective moral values, fail. If no

---

[93] Bennett, J. 1974.

[94] Perhaps a more serious weakness of neutralism is its willingness to allow into the 'moral sphere' a number of (prima facie) non-moral rules (e.g., "Always hold your fork in the left hand"), as long as they are sincerely held and consistently practiced. The same would apply to the pointless rules commanded, say, by superstition. Thus, P. Foot, commenting on a similar position, observes that "the consequences of such assumption are hard to stomach; for it follows that a rule which was admitted by those who obeyed it to be completely pointless could yet be recognized as a moral rule. If people happened to insist that no one should run round trees left handed, or look at hedgehogs in the light of the moon, this might count as a basic moral principle." (Foot, Philippa., 1958, p. 512).

substantive account of morality can be justified, neutralism appears to be the only option left. Yet, there are no reasons to believe that no defensible account of the moral point of view can be given which would encompass the constraints on the form and content of the rules of morality.

*Descriptivism*[95] is the view that argues that any successful definition of morality must include references both to what the moral principles are *about*, and to at least some of the formal features of these principles. Descriptivism here is contrasted with neutralism, which makes no such claims.[96] Descriptivism in its general form is compatible with a number of theories in normative ethics that specify moral rules by reference to something *other* than their actual observance or approval by the moral agents. It implies that simply using the language of morality and employing moral concepts and predicates in description of one's behavior is not sufficient for taking a moral point of view. Rather, certain objective descriptive criteria[97] need to be satisfied to legitimize the use of the moral discourse.

Utilitarianism can serve as a convenient example here, although it is by no means the only ethical theory that would fulfill the general requirements of descriptivism. According to the rule-utilitarianism,[98] an action-guiding principle can be considered as a candidate for the set of all moral rules if and only if the general following of this principle will have a non-

---

[95] The terminology is adopted from Kai Nielsen (1989).

[96] An analogues distinction is sometimes drawn between the two categories of *moral terms*. The so-called 'thin' moral terms (e.g., 'good', 'right', 'ought') say little or nothing specific about the agent, action or the state of affairs that is being so appraised. On the other hand, the 'thick' moral terms like 'brave', 'generous', 'merciful' are at least partially descriptive (in addition to being evaluative). (Cf. Hare, R. M., 1952; Williams, B. 1985, p. 143-5; and McGinn, Colin, 1997, pp. 92-3).

[97] That is, criteria that are relatively stable and do not shift from individual to individual.

[98] Arguably, the act-utilitarianists will recognize the value of the general principles as well, at least as 'rules of thumb' to be used in everyday situations.

trivial influence on the welfare (or happiness) of those affected by these actions.[99] This requirement clearly limits the set of all objects, actions or states of affairs that a moral rule can be about, i.e., it sets the constraint on the possible content of a moral rule. "Always tie your left shoe first," and a number of similar trivial (albeit innocuous) maxims will be placed beyond the sphere of morality, since the general following of these maxims will have (we may safely assume) no connection whatsoever with the woe and weal of human beings.

Moreover, any version of utilitarianism will require that the rules be formulated in a certain way. The formulation of a rule which makes it applicable to specific individuals or a certain group of people only, will be deemed unacceptable, since it fails to do justice to the *impartiality* requirement. In addition, it is often argued, for instance, that *prescriptivity* and *universalizability* are other important formal constraints on moral principles. Other candidates are *publicity* and *practicality* of moral principles (where utilitarianism, it is often argued, falls short). What exactly these requirements imply and how much actual work they can do in sorting out the moral from the non-moral rules is not important at this point. What is significant for descriptivism is that these or other features of principle formation be taken into account by any ethical theory, although *qua* descriptivism it is not committed to any specific set of these features.

In what follows I shall argue that a defensible account of the nature of morality (which will, among other things, provide an acceptable solution to the demarcation problem) should assume the (broadly defined) *empirical* method of investigation, and specify the essential features of the moral point of view along the descriptivist lines. The thesis I shall defend concurs with the strong intuition shared by many people that (in R. Joyce's words), "a

[99] If the effect is positive (i.e., the general following of this rule will create, *ceteris paribus*, more pleasure than pain) the rule is a moral one (in the narrow sense of 'moral'), if it is negative – an immoral one. But in both cases, they are part of the moral sphere (in the wider sense of 'moral').

system of values in which there was no place for condemning Nazi actions simply would not count as a *moral* system."[100] Once a plausible descriptivist definition of morality is formulated as a limiting condition to all normative systems, and once the distinctive features of the moral principles and moral judgments are specified, there will be no need for a separate refutation of neutralism. The acceptability of the descriptivist premise *a fortiori* entails the dismissal of neutralism.[101]

It is worth reiterating that we are *not* here seeking the criterion for a *moral* principle or action as opposed to an *immoral* one, but rather trying to delineate the moral domain proper. The question is what makes a choice, action or a formal rule morally *relevant*, even if only relevant in the sense of being contrary to the requirements of a given system of morality. This ambiguity of the word 'moral' needs to be kept in mind in order to avoid common confusions. In the wider sense of 'moral' it is opposed to 'non-moral', while in the narrow sense it is contrasted with the 'immoral'. Thus a principle "Always consider the interests of the members of the group X as being more important than the interests of all others" is (apparently) a moral principle in the wider sense (in the sense of being morally relevant), but is perhaps an immoral one in the narrow sense of the term.

Given this distinction, it will be clear why some of the well-known tests for the moral status of an action or a principle, such as Kant's requirement for the universalizability of maxims (as 'subjective principles of volition') are not helpful in this context. According to Kant's initial formulation of the categorical imperative, an action is morally wrong if and only if one cannot universalize (i.e., without being inconsistent) the maxim associated with

[100] Joyce, R. 2002, p. 43.

[101] For the specific criticisms of neutralism see Milo, *Immorality*, Chapters 6 & 7.

this action.[102] This decision procedure may at least in some cases help to identify the immoral actions (such as giving a false promise), but it would hardly be useful in sorting out the morally relevant maxims from the rest. Presumably, the maxim "Never look at the full moon after midnight" can be universalized in the required sense without contradiction, but this will not yet explain why it is still a non-moral guide.[103]

In the same way, the basic utilitarian criterion of a moral action (the 'greatest happiness' principle) cannot be used for the demarcation purposes, since it already presupposes a certain conception of the moral sphere – namely, that an action pertains to morality as long as it has an effect on the well-being or satisfaction of interests of at least some sentient beings. In doing so, utilitarianism perhaps presents a plausible material constraint on the moral domain, but this condition itself cannot be obtained or justified by applying a utilitarian decision procedure (which only works *within* morality). By claiming that the morally relevant consequences of an action have to do with the increase or decrease in the 'happiness' of sentient beings affected by the action (rather than, say, with the degree of the overall logical consistency in the universe), the utilitarian has already made a decision as to where the line between the moral and the non-moral should be drawn.[104]

These points about the Kantian and the Utilitarian ethical theories suggest that the question of demarcation between the moral and non-moral is more fundamental than the

---

[102] Kant, Immanuel., 2001, p. 16.

[103] Even if it can be considered a morally *permissible* type of action. This problem is sometimes called the 'problem of relevant maxims.' (Cf. Timmons, Mark. 2002, p. 123). It should also be mentioned that Kant's other formulation of the categorical imperative ('treating as ends, never as means') will be more helpful in determining the morally relevant features of an action or a principle, since it assumes that an effect of an action (or omission) on human autonomy is (in part) what determines its moral status.

[104] I shall argue below that the effect on human welfare *is* indeed a feature of an action that makes it part of the moral domain (morally relevant). But this conclusion will be supported independently by the observation of the common conception of the morally relevant reasons in ordinary moral discourse.

question about the distinction between the moral and the immoral, since the latter issue can only be addressed after the sphere of the moral (in the wider sense) is delineated. The fact that the difference between the moral and non-moral is often assumed to be obvious should not obscure its importance. But, of course, the difference is *not* obvious in many cases. It is sometimes brought against the unsophisticated versions of utilitarianism, that the straightforward application of the 'utility principle' would make the rules like "Everyone ought to improve their sexual technique" into a universal moral obligation. The usual objection then is not that this would be an (intuitively) immoral principle, but rather that it appears to address issues that are outside the moral sphere altogether. The force of this objection depends, once again, on the kind of answer given to the demarcation problem.

In the following section I shall present and defend William Frankena's formulation of the necessary and sufficient conditions for taking a moral vintage point which unambiguously sets apart the *type* of prescriptions, prohibitions and the various 'rules of thumb' that cannot be considered as part of some alternative 'moral code' but should rather be relegated to the non-moral domain. The analysis will focus on the notion of a moral *judgment*. This will, in turn, make the notion of amoralism more concrete: One takes the *amoral* point of view (and thus can be classified as amoralist) if the behavior he exhibits in the *morally demanding* situations is causally based on the non-moral judgments or maxims, where a 'morally demanding' situation is initially defined as a situation requiring a specific moral posture. As a corollary, a distinctly moral action is any action that is motivated by a moral judgment. The demarcation of the morally demanding situations and circumstances from the morally neutral ones will be a special case of the general problem of demarcation.

2. William Frankena's Definition of Morality

William Frankena offers a characterization of morality that builds upon the basic intuitions and empirical observations about ordinary moral principles, but he develops them with much greater precision. It will be helpful to quote Frankena's account of the moral point of view at some length here first:

> My own position, then, is that one is taking the moral point of view if and only if (a) one is making normative judgments about actions, desires, dispositions, intentions, motives, persons, or traits of character; (b) one is willing to universalize one's judgments; (c) one reasons for one's judgments consists of facts about what the things judged do to the lives of sentient beings in terms of promoting or distributing non-moral good and evil; and (d) when the judgment is about oneself or one's own actions, one's reasons include such facts about what one's own actions and dispositions do to the lives of other sentient beings as such, if others are affected. One has a morality or moral action-guide only if and insofar as one makes normative judgments from this point of view and is guided by them.[105]

Frankena lists here four distinct characteristics of the moral domain, every one of which we should examine separately. The very last sentence in the quote above suggests that he takes each of these four features to be necessary and jointly sufficient conditions for taking the moral point of view.[106] We shall see that accepting such a precise specification of the moral domain will not yet commit one to any particular theory in normative ethics. It is likely that a number of different (and, perhaps, conflicting) ethical theories will satisfy the four requirements. This fact will be relevant in the sense that it will tend to confirm that we are dealing here with the theories of *morality*, but it will yield no additional support for any of the specific accounts of the nature of moral obligation and moral rightness.

---

[105] Frankena, W., 1963, pp. 113-14.

[106] The last, fourth, requirement, of course, is only conditionally necessary – it applies only to cases when the self-referential evaluative judgments are made.

2.1 The Proper Subject and Object of a Moral Judgment

The initial feature deals with the possible objects of evaluative judgments. Frankena lists actions, desires, dispositions, intentions, motives, persons, or traits of character as the only acceptable objects of a *moral* normative pronouncement. First, the most basic constraint here is the constraint on the domain of the possible moral subjects. On one natural reading of Frankena's position, it is only *human* actions, desires and traits of character that can be evaluated from the moral point of view. We do not need to go here into explanation of this unique status of human beings (although, an adequate explanation can perhaps be given along the Kantian lines). But it is important to note that this proposed inclusion of the set of all moral agents into the set of all human beings[107] is not a self-evident statement. There have been claims to at least two types of exceptions, which have to be addressed briefly. First, Jonathan Glover in his book "Responsibility" brings evidence that during the Middle Ages it was not uncommon that animals, not only humans, were in the role of the criminally charged defendants during the legal trials,[108] and it was often assumed that their allegedly harmful actions and evil intentions can be legitimately evaluated from the moral and legal perspectives.[109]

Secondly, there is a longstanding tradition that ascribes moral properties and accepts moral judgments about God and a number of other spiritual beings that can have specific goals and exhibit purposeful behavior. The alleged fact of their rationality is often seen as the

[107] Inclusion does not imply coextensiveness. Only human beings can be the subjects of a moral judgment, but it is not true that all human beings are moral subjects in this sense (e.g., very young children or mentally retarded people are exceptions).

[108] Glover, Jonathan. 1970.

[109] By dismissing such attitude about animals as obviously preposterous, I am not suggesting, of course, that *human* actions towards animals cannot be morally relevant.

feature that immediately qualifies them for membership in the moral domain. Thus Kant in the *Preface* to the "Foundations of Metaphysics of Morals" extends the scope of the moral law in order to cover all possible rational beings, regardless of their biological peculiarities:

> Everyone must grant that a law, if it is to hold morally, that is, as a ground of an obligation, must carry with it absolute necessity; that, for example, the command "thou shalt not lie" does not hold only for human beings, as if other rational beings did not have to heed it, and so with all other moral laws properly so called; that, therefore, the ground of obligation here must not be sought in the nature of human being … [but] a priori simply in concepts of pure reason; and that any other precept …, insofar as it rests in the least part on empirical grounds, can indeed be called a practical rule but never a moral law.[110]

Indeed, Kant takes it to be a distinctive and essential feature of a moral law that its scope is not limited by any biological contingencies. The scope of a moral prescription, for Kant, is *universal*, whereas the scope of any non-moral rule (e.g., social conventions) is necessarily limited by the empirical circumstances of the occasion where the rule is put forward. To claim that a law necessarily applies to *all* rational beings (including, perhaps, the non-human spiritual creatures) implies that one cannot ground this law in any contingent fact about humans (e.g., any desires that humans have, even if these desires are shared by all members of the species, as the desire for happiness is said to be). The universal application of the law to all *rational* beings requires grounding of this law in the notion of rationality itself.

Regarding the first alleged exception, I shall assume that the medieval practice of holding animals morally responsible requires no special refutation. Since I take self-awareness and the ability of abstract reasoning to be necessary conditions for being a competent member of the moral domain, whatever level of rationality some higher animals exhibit, it falls short of the level required for being a genuine moral agent. But the ability to

[110] Kant, Immanuel., 2001, p. 2-3.

reason is not *sufficient* for that. Kant clearly overemphasizes rationality as being the only ground of morality. There is something else which is just as indispensable for any moral agent *qua* moral agent. It is hard to capture this feature by a single term, but perhaps a descriptive phase 'capability to be hurt' will do the job. It seems that the fact of our human fragility is what lies at the foundation of most systems of morality. We are standing in a moral relation to other humans precisely because they *can* be benefited or harmed by our actions, and vice versa.[111] Imagining a rational being who cannot be adversely affected by any actions of other agents (as God is said to be) stretches our ordinary moral intuitions to the point where we can no longer recognize him as a being about whom a moral judgment can be made or whose actions can be morally evaluated.

This, however, leaves open a possibility that other non-human, but rational *and* fragile creatures are part of the moral order. As a theoretical possibility, I think, the existence of such creatures should be allowed. But since we have no credible evidence to their actual existence, we can safely assume (for now) that the members of our own species are the only moral agents in the universe.

After indicating the proper subjects of moral judgments and requirements (i.e., ourselves), we can concentrate on the morally relevant features of our nature and our behavior. Frankena argues that we can morally evaluate one's actions, desires and dispositions, but we are not making a *moral* judgment when we, for instance, are making normative comments about one's looks, ethnic origin or one's linguistic accent.[112] There are

---

[111] Cf. Rousseau: "Man's weakness makes him sociable. Our common sufferings draw our hearts to our fellow-creatures; we should have no duties to mankind if we were not men." (*Emile*, p. 182.)

[112] This seems to be partly because the features cited lie outside the domain of our freedom. One does not *choose* to be born at a certain place, with certain biological features, etc.

intuitively clear constraints on what can be taken as a morally relevant feature of agents. Not every aspect of a person can be the proper bearer of moral value. Frankena's extended list of the morally relevant features can perhaps be reduced to *intentional actions* as the primary bearers of moral value.[113] On the other hand, one's intentions, desires, motives and character can be taken as morally relevant in a *derivative* sense only – their moral relevance is parasitic on the fact that they tend to affect and produce certain types of actions.[114] In a possible world where there is no connection between one's 'inner' state and one's outward behavior, the moral evaluation of one's intentions and desires would be quite out of place.

If actions are the primary bearers of moral value[115], one is taking a moral point of view if (minimally) he makes a normative judgment about one's actions or any other feature of inner state or character that may affect or lead to actions. At the same time, a refusal to make such a judgment in relevant circumstances may (occasionally) indicate a non-moral stance assumed by an observer. But clearly not every instance of a judgment about human behavior warrants the name of a moral judgment. I may strongly disapprove of one's conduct without yet expressing a moral disapproval, as when I disapprove of the handling of the puck by a hockey player on the ice. In addition, Frankena argues, one must be willing to universalize one's judgment and be able to justify it in the right way.

---

[113] The question what precisely makes *some* actions morally relevant events cannot be fully discussed here. I suggest that there are two interrelated features. First, an action is morally relevant if it is intentional, i.e., *free* in some acceptable sense of this word. Secondly, it is morally relevant if it is a type of event that may affect the wellbeing of others.

[114] Kant famously argues for what appears to be the opposite view. For Kant, it is the will (which includes intention) that is good or bad in the primary sense, and actions have positive moral value only derivatively, in so far as they are the products of the good will. (Kant, I. 2001, p. 1) But Kant still recognizes the connection between good will and moral actions when he argues that the moral status of an action might be questionable *because* it might be the result of, say, selfish motives. His further claim that a good will would still "shine like a jewel" even if no actions were to follow seems counterintuitive.

[115] We do sometimes apply moral predicates to states of affairs, e.g., "this system is unfair.", but these statements can in principle be analyzed as referring to the actions of particular persons or groups of people.

2.2 The Universalization Condition

The requirement of universalizability is a formal requirement on a moral judgment or principle, a constraint that deals both with the way it is (ideally) formulated and its intended scope. It is a requirement that excludes from the moral realm all evaluative formulations whose validity is essentially tied to the specifics of the place, time and circumstances in which a judgment is made. If one judges a certain action to be morally right or wrong, one is committed, given this requirement, to pass the same judgment on all relevantly similar actions as well. The idea is intuitively clear, but the exact interpretation of this condition will have to be worked out.

A brief caveat first. Frankena's universalizability constraint must be clearly distinguished from Kant's universalizability test for maxims. When Kant argues that the (objective) moral status of a maxim depends on its being universalizable, he makes no reference to one's willingness or unwillingness to take this test. Whether the maxim can or cannot be universalized without contradiction is an objective fact, which, presumably, does not depend on someone actually applying the procedure. Frankena, on the other hand, talks about the '*willingness* to universalize one's judgment' and says nothing about the actual success or failure of this process. This verbal difference is explained by the differences in goals. As it was mentioned before, Kant's test was meant to distinguish between the morally permissible and morally prohibited actions, whereas Frankena here is concerned with identifying the morally *relevant* judgments about actions, regardless of the actual moral status of those actions. Thus Kant's 'decision procedure' introduced in his first formulation of the Categorical Imperative is irrelevant to the present discussion.

What does it mean then to be willing to universalize one's evaluative judgment? If universalization is a matter of degree, to what extent should one be willing to universalize his judgment in order to assume a moral standpoint? J. L Mackie discusses the three distinct stages of universalization process, where each successive stage presupposes ever greater abstraction from the particular features of an act and circumstances that prompted the judgment and a person who makes it.[116] On the first stage we rule out mere numeric differences between the two actions or persons as irrelevant. More formally, this implies that "a judgment containing a proper name or indexical term used not as a variable but as a constant will not yet be universalized,"[117] and thus anyone making this kind of judgment is not making a moral judgment, but rather a non-moral one. In this sense, the requirement of universalizability amounts to the requirement of being sincerely willing to apply the same prescription or evaluation to oneself as well as to others, as long as there are no relevant qualitative distinctions between the cases. This kind of universalizability, as Mackie notices, rules out at least one variety of egoism, viz., the one which claims that it is right, morally speaking, *only for me* to pursue my own interests and I should have a regard for the interests of other only insofar as it contributes to my own interest.[118] This thesis of *personal* egoism (as Kai Nielsen calls it) fails to satisfy the minimal universalizability requirement, since it denies that the same principle ought to be binding on other agents without showing, however, that there is a relevant difference between the cases. Thus anyone adopting it as an action-guide will be assuming a non-moral vintage point. The amoralism of this type of ethical

[116] Mackie, J. L., 1977, pp. 83ff.

[117] *Ibid*., p. 84.

[118] It does not rule out, however, other varieties of ethical egoism which make it into a universal prescription that *everyone* ought to seek their own good.

egoist follows from his failure to satisfy one of the necessary conditions for having a morality.[119]

The second and the third stages of universalization are operating on the higher levels of abstraction. On the second stage, in making a decision as to whether the maxim is universalizable, one is required to imagine oneself in another man's place, in his specific conditions, and ask whether he would still will this maxim in those circumstances. The proposed 'thought experiment' where one imagines himself to be in the place of a less financially successful, or a less healthy person would exclude, it is argued, a number of unfair principles which will otherwise pass the first stage of universalization.[120]

The last (third) stage of universalization involves even deeper appreciation of the differences between the moral agents. In addition to considering the differences in external situations and abilities (e.g., wealth, social status) between myself and another person, it requires taking on his values, preferences, desires and ideals.[121] What sort of principles would I endorse had I been not only in his situation, but also with his set of values? As

---

[119] "Personal ethical egoism isn't a malign, satanic or cynical morality because it isn't a morality at all. We might say of some, from our-point –of-view, 'far out' moral code – say Aztec's or the Keroki's moral code – that it was a perverse or even an immoral moral code. By contrast, personal egoism is a mere contempt for moral considerations altogether and this is not a moral code at all. It is not even something that could be intelligibly and coherently proclaimed as a morality." (Nielsen, Kai, 1989, pp. 158-9).

[120] E. g., the financially secure and relatively healthy person may sincerely and consistently endorse the principle that one should always pay for his or her medical needs.

[121] There exists an interesting parallel between Mackie's three stages of universalization and the three levels of moral development from Lawrence Kohlberg's 'stage theory'. Kohlberg (1970) argues that the six basic stages of moral development can be demarcated into three levels: preconventional, conventional, and postconventional (principled). The three levels can be thought of as three way of relating the self to moral expectations of society. In the preconventional level the universalization of rules is done mechanically, barring all exceptions. In the level of conventional morality the awareness of the motives, intentions, or moral sentiments of others allows for limiting deviations from the universalizable norm. And the final, postconventional (principled) level requires an even deeper appreciation of the reality of other - an ideal role-taking. On this level, the need to accommodate the perspectives of each party in a complex social situation is taken in account when formulating the relevant moral rules. (Lapsley, 1996, pp. 70-3). Yet, even the first level of a moral development yields genuine moral judgments, and a person (a child) at this level is undoubtedly taking a moral perspective.

Mackie observes, at this last stage of universalization, it hardly makes sense to talk of putting *oneself* in someone else's place, "since hardly any of oneself is retained."[122]

What degree of universalization of one's judgments is minimally required, then, in order to assume a moral point of view? Mackie argues that only the first stage of universalization is relevant for the demarcation problem. The last two stages from Mackie's account, which taken together are not unlike Rawls' requirements for the selection process behind a 'veil of ignorance',[123] may potentially be employed for identifying the moral principles (as opposed to immoral ones),[124] but one's willingness to universalize his judgments about actions, intentions, desires or traits of character along the lines of the first stage is (*ceteris paribus*)[125] sufficient for that person to assume a moral (as opposed to a non-moral) point of view. There will still remain vexing problems in specifying when the two or more situations are relevantly similar, or what counts as a morally irrelevant difference between the two persons, but anything more demanding than the first stage of universalization will threaten the very existence of the moral point of view.[126]

---

[122] *Ibid*., p. 93.

[123] Rawls, John. 1971.

[124] Although Mackie himself admits that "it is doubtful that any principles will pass so sever a test." (p. 93).

[125] I. e., assuming that the last two conditions from Frankena's list are met.

[126] Mackie is rightly concerned that if the more demanding interpretation is accepted, and it turns out that no one ever is willing to put his judgments to the rigorous test of, say, the third stage (ideal role-taking), it would follow that no one is ever making a moral judgment. The difficulty in attaining the higher stages of universalizations is nicely illustrated by the following interpretation of the Golden Rule by a ten-year-old: "Well, it's like your brain has to leave your head and go into the other guy's head and then come back into your head but you still see it like it was in the other guy's head and then you decide that way" (quoted in Kohlberg, 1970, p. 642). Moreover, adopting anything more demanding than the first stage would permanently exclude autistic individuals (who are inherently unable to look at the world from the other person's perspective) from the moral domain. Yet, as Kennett argues (2002), many able autistic persons are capable to assume a moral point of view by following the formal process of unbiased rule-application.

I believe we should agree on this point with Mackie's judgment.[127] As far as the universalization condition is concerned, it seems to be sufficient for taking a moral perspective that a person be willing to apply his normative judgment to anyone or anything that is relevantly similar to the instance in question. One might argue that there is something rigoristic in the requirement of universalization along the lines of the first stage, since it is only concerned with the unbiased application of the principle or judgment, but hardly with its *content*. And there is something paradoxical in saying that the content of the principle does not matter (for identification of the principle as a moral one) as long as the formal condition of universal application is satisfied. But all this shows is that universalization condition *alone* is not sufficient for fully specifying the moral point of view. It is rather a purely formal condition that must be further supplied by the constraints on the possible justification of the moral principles and judgments.

2.3 The Justification Condition

The third condition from the Frankena's account deals with the kind of *justification* that a proponent can provide for his normative judgment, granted that the judgment is about human actions and the proponent is willing to universalize it in the relevant sense. The justification condition is the central issue here. The first two requirements are still not restrictive enough to eliminate a number of *prima facie* non-moral rules and non-moral evaluative judgments, such as "it is despicable that Glenn does not cross his hands twice when entering a house." The judgment is about Glenn's actions, and the proper name used in this context is a variable that one might be willing to substitute with a universal quantifier. But it is not a moral judgment unless one makes it for the right kind of *reasons*:

[127] But perhaps *only* on this point.

> It seems to me that what makes some normative judgments moral, some aesthetic, and some prudential is the fact that different points of view are taken in the three cases, and that the point of view taken is indicated by the kinds of reasons that are given.[128]

The right kinds of reasons, according to Frankena, are "facts about what the things judged do to the lives of sentient beings in terms of promoting or distributing non-moral good and evil."[129] The non-moral good and evil can be variously specified, but it is traditionally explained in terms of pleasure and pain. The less tendentious formulation will employ the notions of well-being and ill-being, since the descriptive element contained in these two terms is marginal (as compared with 'pain' and 'pleasure' or even 'happiness' and 'unhappiness'), which in turn allows for greater flexibility in formulating a *substantive* theory in normative ethics.[130] Furthermore, Frankena specifies the possible recipient of the non-moral good as a 'sentient being', a category that encompasses humans, animals and (potentially) any other creatures capable of experiencing pain and pleasure.

Frankena offers the third condition as an important factor for distinguishing between moral and non-moral *judgments*. One makes a moral evaluation in so far as he justifies his judgment by reference to the facts about what "the things judged" do to the well-being of sentient creatures. We have seen that "the things judged" in this context are primarily human intentional actions, and derivatively, the inner states of human agents. Willingness to point to these facts as reasons for one's judgment is a mark of a moral attitude. As a further

---

[128] Frankena, W., 1963, p. 110.

[129] *Ibid*., p. 114.

[130] Hare objects that "to put this restriction upon the use of the term 'moral' is to write some kind of utilitarianism into its definition." (1963, p. 163). This seems to be an overstatement, since all that such content-based definition of morality implies is that the fact about human welfare must constitute at least a *relevant* reason for making a moral judgment. There is nothing in this definition of morality that specifies the weight one should assign to these reasons, or the extent of impartiality with which one should consider these facts. Both Kantian ethics and Utilitarianism are compatible with Frankena's understanding of morality defended here.

consequence, we can say that those actions that have this external property of affecting the distribution of well-being among other sentient beings are morally relevant actions, and the states of affairs that involve these actions as part of their description are morally relevant *situations*. On one plausible interpretation, a morally relevant action *results* from a morally relevant judgment,[131] even in those cases when the corresponding judgment is not formulated explicitly before the action takes place. In Frankena's own phrase, in acting one is in important sense "guided by [evaluative judgments]."[132] Thus, many of the same criteria should apply to both types of events: intentional actions and normative judgments.

There is, however, an important ambiguity in Frankena's characterization of the third condition. This ambiguity can be illustrated by the following two examples. In the first case, a witch-doctor makes an evaluative judgment: "It is always wrong to spit in the direction of the Great Forest", and, when prompted, explains it by reference to the bad-tempered Evil Spirit, who lives in the forest, and who tends to get offended by these kind of disrespectful behavior. As a result, the Spirit might destroy the crops, which will hurt the community. In the second case, a missionary claims that it is wrong to divert the river from its current course, and justifies his judgment by arguing that the lack of water will cause a drought, which will hurt the crops, and, ultimately, hurt the community. The relevant difference between the two examples is that in the first case the witch-doctor is clearly mistaken in his factual beliefs about the damaging effects of the spitting, whereas in the second case, the missionary is probably correct about his predictions. Does the witch-doctor then make a genuine *moral* judgment?

---

[131] See Audi, R. (2006) for a defense of this view.

[132] *Ibid*.

More specifically, we can ask whether the justification story in terms of welfare changes for one's evaluative judgment has to be also *accurate*, in addition to it simply containing references to (sincere) *beliefs* about the alleged consequences of actions, in order for that judgment to qualify as a moral one.[133] The emphasis on the objective accuracy of the justificatory story would exclude from the set of all moral judgments those prescriptions and value judgments which are based on clearly mistaken factual premises; whereas the more liberal interpretation would recognize the prescriptive claims of the above-mentioned witch-doctor as moral pronouncements despite the falsity of the justification story. Before settling the issue, we should look at both alternatives more closely.

Ronald Milo, among others, embraces the latter interpretation of the third condition when he says:

> If one holds that it is wrong to step on the lines of a pavement and if one's reason for thinking this is that this will cause great harm by bringing on a plague, then one may be said to accept this principle as a moral one.[134]

On Milo's view the only thing that matters in deciding whether a person assumes a moral point of view or some other non-moral normative perspective is the subjective state of the person who makes a value judgment of some kind. In so far as he (sincerely) *believes* that stepping on the lines of the pavement will bring great harm, his proscription to do so assumes a moral status, and no further connection of those beliefs with reality is necessary.

This interpretation, however, cannot be accepted for the following reason. The emphasis on the subjective side of the justification of one's normative judgments threatens the descriptivist account of morality to collapse back into neutralism. If one's beliefs about

---

[133] This question is not identical to the more familiar question of what should determine the *moral worth* of an action, its actual consequences or the intention of the agent.

[134] Milo, R. 1984, p. 197.

the facts how "the things judged" (e.g., actions) affect the welfare of sentient beings do not have to be true or even probable, but only, perhaps, sincere at most, then *any* normative judgment can be turned into a moral one by acquiring a required set of (false) factual beliefs. It is conceivable that the normative claim "All members of the ethnic group *X* are evil, and thus should be exterminated" can be universalized in the required sense (even including the willingness to apply this prescription to the one who utters the claim if he turns out to *be* the member of this group), and justified by references to the alleged immense harm that the members of this group produce for the rest of the world. As long as the factual story is sincerely believed, it will be a moral judgment, on Milo's interpretation. Yet, part of the motivation for embracing the descriptivist account of the nature of morality in the first place was precisely the promise that an account alternative to neutralism would exclude from the moral domain any system of action-guiding rules which incorporates this or similar prescriptions (e.g., 'Nazi morality'), and would further specify objective criteria for such exclusion. Hence, if a descriptivist account is to be a genuine alternative to neutralism, a more restrictive interpretation of the third condition needs to be adopted.

I suggest though that it is not necessary to demand that the justification of one's normative judgment in terms of changes in welfare of sentient beings as a result of intentional actions should *actually* be true on all occasions when a judgment is made. Even if the missionary is accidentally mistaken in this particular case in his estimation of the adverse effect of the diversion of the course of the river, he is still making a genuine moral judgment, since it is presumably more likely than not that substantial harm will be done as the result of these manipulations with nature. It seems sufficient to accept a probabilistic account of changes in welfare, using general experience and available scientific information as our

guides in specific situation. To be sure, all initial references to probability are inherently vague so that a number of borderline cases will always remain (e.g., *how* probable the alleged future harm needs to be in order for a prescription to qualify as a moral judgment, etc.). But adopting this interpretation will set certain constraints, and will at least allow us to say this much: Milo's example of "it is wrong to step on the lines of a pavement" will be a non-moral judgment in *most* contexts, whereas "it is wrong to let children play on thin ice" will be a *prima facie* moral claim, since the harm ensuing from playing on thin ice is much more likely to follow than the alleged plague that is believed to occur as a result of stepping on the lines of a pavement.

Finally, Frankena's last condition has a limited application since it concerns only judgments about oneself or one's own actions. In those cases one is making a *moral* judgment, he argues, when "one's reasons include such facts about what one's own actions and dispositions do to the lives of other sentient beings as such, if others are affected."[135] This fourth condition is best seen as a special case of the previous one, since it concerns the kind of justification that one can provide for his self-referential evaluative judgments or actions. Once again, the acceptable justification will refer to the effect that one's actions have on the well-being of others. The phrase "*other* sentient beings" and the appearance of the hypothetical clause in the last line, viz., "if *others* are affected", suggests that Frankena does not consider the effect on the well-being of the agent himself as the relevant factor here. If doing X will harm the agent, and only the agent himself, and, as a result, the agent makes the following judgment: "I was wrong in doing X", he is not, according to this reading, making a moral judgment proper, but rather an evaluative comment from the prudential point of view.

[135] Frankena, W. 1963, p. 114.

In other words, on this view one does not have specific moral duties to oneself, unless it also affects the well-being of others.

My interpretation of Frankena's characterization of morality presented and elucidated in this section will provide a necessary foundation for the discussions to follow. In particular, I have argued that there are at least three essential features constitutive of the moral point of view, which deal with the content, the form and the justification of a normative judgment. I have suggested that human actions (those of others and possibly one's own) are the primary targets of a moral attitude expressed by means of an evaluative judgment. In addition, the person must be willing to universalize his judgment along the requirements of the first stage of universalization, and should be able to present a justification which will include references to how the actions judged affect the well-being of sentient creatures. But the alleged effect on the well-being of sentient creatures need *not* be true of the particular actions, as long as the arguer accepts a more general principle which mentions a *type* of actions with such an effect, and as long as the arguer *believes* that the particular action judged is an instance of that general type. Lastly, the features constitutive of the moral point of view presented in this section correspond to and can be derived from the ordinary moral intuitions.

A *descriptivist* account of morality implies that certain attitudes, normative judgments, principles, action-guiding systems or particular actions can be *excluded* from the moral realm using the definite criteria set up by the account. This account can be characterized as a kind of *moral naturalism*, since it holds that certain factual considerations necessarily count for or against the ascription of moral terms.[136] On the other hand, it does

[136] This is not the place for the detailed discussion of the traditional objections to naturalism. For a thorough discussion of various objections including the criticism of the 'naturalistic fallacy' charge see, for instance, Simpson, Peter. (1987) *Goodness and Nature: A Defense of Ethical Naturalism*, and Kohlberg, Lawrence.

not yet commit us to any substantive moral view (e.g., utilitarianism, or deontlogy), and is quite compatible with a number of different theories in normative ethics.

3. The Normative Aspect of Amoralism

As yet no argument has been advanced to show that *having* morality (i.e., assuming a moral point of view) is not *optional* as, for instance, having political views might be. One is not required to take a definite stand on a political issue, but the situation is (presumably) entirely different when one finds himself in a morally demanding situation. Witnessing a brutal murder calls for a *moral* judgment and an appropriate action, and a failure to heed the call is considered to be a much more serious lapse than, say, a failure to adopt a certain attitude toward a political party's proposal on increased taxes for farmers (unless the issue has clear moral implications). We may describe a person who is completely indifferent toward the political campaigns, elections and programs as apolitical, but the connotations of this term are by far less negative than the ones we associate with the term 'amoral'. 'Amoralist' is partly a descriptive term that refers to a person who is indifferent to the requirements of morality, but it also has rich evaluative meaning. When we label someone an amoralist, we use the term primarily as a judgmental tool, rather than a purely descriptive device. What we intend to say is that there is something *wrong* with being amoral, it is a condition and a way of life that is not *normal*, and the person designated by this term is clearly a *bad* human being *qua* human being. Furthermore, this normative layer seems to be much more conspicuous in ordinary speech than the purely descriptive, cognitive meaning of the term.

---

(1971) *From Is to Ought: How to Commit the Naturalistic Fallacy and Get Away With It*. M. Devitt (2002) presents an outline of a plausible version of moral naturalism.

The irreducible normative aspect of the phenomenon of amoralism puts new demands on the present explanatory project. An explanation is an answer to a 'why' or a 'how' question, and in the given context we want to know why some people exhibit the traits of character and behavior which we have previously identified as amoral. But given the peculiar nature of the explanandum, the explanans, it seems, cannot be limited to the list of causally relevant facts and the neutral description of the antecedent conditions, i.e., the kind of explanatory account one would expect, say, in reading why animals have no language or why some people have darker skin. A full account of amoralism would not only identify the features that are causally responsible for the condition, but would also address the question why one *ought not* to be in this condition, i.e., why it is an unfortunate, undesirable and a terrible condition to be in.

This last point brings with it a whole baggage of traditional philosophical problems, all of which converge on a single question: "Why be moral?" If amoralism is an *abnormal* human condition, we must first define the *normal* condition of human beings, and present an extended argument to the effect that everyone, in some important sense, *ought to* prefer the normal condition to the abnormal one, where the power of this 'ought' does not rest simply on semantic distinctions (i.e., it is not *a priori* true).

Historically, in response to this challenge at least *two* different directions have been explored. I shall now briefly summarize each of them in turn. First, one may refer to some actual and allegedly universal human interest (e.g., happiness or pleasure), and argue that, as a matter of empirical fact, amoralism as a way of life and thought is detrimental to fulfilling this interest.[137] We may call it the 'common interest' theory. The explanation of amoralism in

[137] Plato is the classical source here, but similar arguments are proposed by Hobbes, Hume, J. S. Mill, et al.

the context of this theory will then proceed by identifying some *mistake in reasoning* of an amoralist, since he cannot now *rationally* want to be amoral given the operative goal of happiness and the truth of the empirical premise.[138]

This appears to be an initially promising strategy, albeit with important limitations. The main limitation is the inherent ambiguity of 'happiness' (or any other plausible substitute). An attempt to specify the term by listing the conditions of happiness is likely to falsify the empirical claim that everyone wants happiness in *this* sense (e.g., as family life or intellectual activity); whereas leaving the term undefined makes the universal ascription of a desire to be happy empty and uninformative. If practically anything can be included under the denotation of 'happiness' (i.e., '*whatever* makes one happy'), the philosophically interesting claim "Everyone wants to be happy" is reduced to a trivially true statement "Everyone wants something." The truth of the last claim is, for the obvious reasons, unhelpful in constructing a case for the superior status of a moral life.

The second direction requires more assumptions and a more rigorous argument. In essence, this line of reasoning appeals to the notion of a *human nature*, and treats (true) morality as being in accord with this nature, and amoralism (and immoralism) as being unnatural and perverse. In this tradition, the specific human essence is a constant that is shared by all members of the species regardless of one's cultural background or a place in history. Discovering this essence (which is usually sought in non-biological aspects of human existence) would allow one, in principle, to work out a pattern of behavior and a life-style that is conducive toward developing and sustaining this immutable human nature. Everything

[138] I.e., the premise that states that it is a *moral* person who is more likely to be happy, experience more pleasure, satisfaction, etc. The truth of this premise, however, is by no means obvious. See Cahn (2004) and Vitrano (2008) for the argument that a (clever) immoralist might just as well or even more likely achieve happiness.

that serves this goal is thus natural or in accordance with nature, and, on the contrary, the type of actions or attitudes that cannot be seen as natural are for this reason wrong, prohibited and, in general, immoral. The amoralist, on this view, is similar to a sick person, whose condition is abnormal even if he fails to appreciate his abnormality (say, due to the severity of illness).[139] In effect, one ought to be moral simply because one is a human being, and morality is part of what it means to be human.

The main weakness of the 'Human nature' theory is precisely the notion of human nature itself. First, historically there have been intractable disagreements about what exactly is most essential about us (e.g., rational element for Aristotle; spiritual element for religious traditions; productiveness for Marx[140]). Secondly, we may question whether 'fixed' human nature, as traditionally understood, exists at all. The accumulated data from the fields of anthropology, sociology and other social sciences has thoroughly discredited in the eyes of many scholars the ancient conception of the human 'core' that retains its essential features in all historical contexts. Ruth Benedict expresses a common enough opinion when she says: "We do not any longer make the mistake of deriving the morality of our locality and decade directly from the inevitable constitution of human nature."[141] And when J. P. Sartre later proclaims that "there is no human nature, since there is no God to conceive it,"[142] he is merely pointing to the widely recognized (by that time) failures of the classical teleological project, which tried to infer the true religion, morality, and even the optimal political order

---

[139] Cf. Aristotle, *Nicomachean Ethics*, 1145 a15-b05.

[140] See Anthony Quinton's *Has Man an Essence?* (1975) for a comprehensive discussion of the various historical attempts to locate the essence of a human being.

[141] Benedict, R. 1989, p. 47.

[142] Sartre, 1957, p. 11.

from the metaphysical concepts of the human *telos* or human essence. The whole project, as Sartre observes, seems to be intrinsically tied with the religious worldview, and becomes increasingly less plausible with the emergence of a sound secular alternative.[143] This is not to say that no version of the 'human nature' theory can be defended in our times,[144] nor can we ignore the possibility of a non-theistic variant of this view. But given the overall direction of modern philosophy the burden of proof that is on the side of the defenders of 'human nature' conception is heavy indeed. The proposal, elaboration and defense of a teleological theory of this kind and its implications for moral philosophy is a substantive project which would require more space and time than can be afforded in the present work. On the other hand, given the sheer amount of critical literature, it is no longer possible simply to *assume* the truth of this theory as a premise in a further argument.

It appears then that neither the 'common interest' theory nor the 'human nature' theory can be easily employed here to set up a conceptual background against which the condition of an amoralist would look unambiguously deviant. However, we have seen that the strong normative connotations of 'amoralism' require furnishing a sense in which we can view an amoral person as *abnormal*, and interpret his characteristic pattern of behavior, his attitudes and emotive states as manifesting a *disorder* of a certain kind.

But perhaps we can sidestep this difficulty and avoid the complex metaphysical issues by resorting to a more modest (in its philosophical aspirations) notion of a *statistical* deviance. This bypassing maneuver assures that we can meaningfully attribute abnormality to

---

[143] It must be observed that teleology may be connected to religion in two different ways. The tradition of the ancient philosophy tends to infer to a religious worldview as a *conclusion* of the teleological project; whereas the modern interpretations (including Sartre's) view religion as an assumed *premise* in all teleological arguments. Only on the *latter* reading would the non-existence of God fundamentally undermine the teleological framework. (This point is due to Prof. Simpson. I am grateful to him for pointing out this difference.)

[144] Cf. Simpson, P., 1987.

a kind of person described in Chapter One on the basis of his observable idiosyncrasy in matters of social relations, as compared with the majority of members in a given population. We do treat people with unusual color perception as abnormal (in their vision), where the *norm* is established by the perception of the majority. On a planet populated by creatures who perceive in inverted spectrum only, the standard of normal vision would be quite different. Likewise, what makes the condition of an amoralist *prima facie* abnormal is the fact that he is evidently in the minority insofar as his interpersonal and social skills are concerned. This is clearly not *all* that is wrong with the amoralists, but it at least suggests *a* sense which initially allows us to apply the categories of normal vs. abnormal to these peculiar individuals.

Admittedly, the reference to statistics captures only a small part of what our ordinary intuitions suggest is true of the amoralist. Not only his condition is abnormal in the sense of being unusual (a genius or a 'moral saint' are equally rare), but we also want to say that there is something inherently evil or, at least, pathological in his amoral state. And the 'evilness' of these people is logically independent from their relative representation in a society. On encountering a person devoid of the usual inner constraints and apparently experiencing none of the morally relevant emotions (e.g., guilt, shame, remorse, empathy), we tend to feel strong aversion against the individual, a reaction which in part can be explained by our realization of how fundamentally *different* he is from the rest of us.[145] This fact is significant. As Williams notes in his discussion of a psychopath (a kind of amoralist), "the psychopath is, in a certain way, important to moral thought; but his importance lies in the fact that he

[145] The reaction of aversion is nearly ubiquitous, singular voices to the contrary notwithstanding. Among the dissenting opinions we may mention F. Nietzsche's glorification of the *Übermensch* (who can be interpreted as amoralist), and, more recently, a book by Alan Harrington, named *Psychopaths* (1972). In this original presentation the author voices the following view: "Although originally founded upon an anti-social condition, it [psychopathy] offers exciting new alternatives to the way we have lived until now…What was formerly diagnosed mental illness has turned into the new spirit of the age…Conceivably the times are calling for an idealized version of the psychopath as savior… [etc.]" (pp. 48-50).

appalls us."[146] But unlike encountering a person with gross physical deformities, or the one possessing a rare artistic genius (in both cases we, the ordinary people, may plead a fundamental difference, and in the former case the initial reaction of disgust is also common), the reaction we exhibit toward the amoralist is rooted in our intuitive sense of the necessity of a moral attitude for a genuinely *human* life.

The amoralist is not only in the minority (which allows the initial identification of his condition as abnormal), but there is, moreover, something (we feel) seriously *wrong* with him. But even if (psychologically speaking) the strength of our intuitive revulsion of amoralists can be *explained* by reference their statistical abnormality, it cannot be *justified* in a similar way. If at the moment neither reference to the 'common interest' nor to 'human nature' is a possibility for us, how else can we justify our judgment about the wrongness of amoralism?

In my view, the only alternative left is to grant the evaluative language that we normally use in our reactions to amoral individuals a type of *autonomy* which does not require any further reduction or analysis. Steven Ross proposes to deconstruct all the traditional approaches to valuation in precisely this way.[147] According to his view, the existing decisive considerations against the various analyses of value, including the teleological projects, do not yet force one to accept the unbridled relativism with respect to evaluative pronouncements (both moral and non-moral). We can safely reject the 'old' theories of values, Ross argues, *and* preserve the following two claims: "(1) certain features of persons

[146] Williams, B., 1972, p.10.

[147] Ross, S., "Evaluating the Emotions" (1984).

are in fact genuinely good or bad, and so certain terms refer to genuinely good or bad qualities; (2) the presence of these features is governed by criteria in fairly clear ways."[148]

The main problem with the traditional approaches to evaluative terms, according to Ross, was the assumption that these terms require further reductive analysis in terms of interest, utility, subjective state of approval or disapproval or else in terms of the 'proper function' of the object being evaluated. This assumption, however, is unwarranted and points to a certain 'category mistake' when the evaluative language is supposed to refer to something above and beyond itself:

> Evaluative language does not pick out a thing, not even an evaluative thing. Rather, we simply have *a way of talking about* certain things in the world, a set of interests, that the world shows itself to answer to. There are evaluative concepts, and these concepts mean *exactly* what they say they do. They do not need to be 'cashed out' in some further description.[149]

Even though, as Ross admits, our ascription of evaluative terms does have some connection with our *interests,* it would be wrong to suppose that we can *define* evaluative pronouncements in terms of our interests. No type-type reduction is possible here. Evaluative language is used primarily to express a certain kind of interest "before the world", but that does not exclude that it may *in addition* be used referentially or descriptively – i.e., to refer to some descriptive features of the world. The application of the term 'cruel' to a person's actions connotes disapproval of these actions as being contrary, say, to our interest in a pain-free existence, but its 'factual content' will vary depending on the context of the utterance

[148] *Ibid*., p. 318.

[149] *Ibid*., pp. 319-20.

(e.g., shooting a convict in the head may be properly evaluated as a merciful action when the only other alternative for him is being burnt at a stake alive).

Ross's analogy with evaluation of art is particularly instructive here. Saying that Queen of the Night's aria in *Der Zauberflöte* is 'expressive' is that "that's the *kind* of accomplishment this aria is and so it brings the kind of involvement it does."[150] There is no further need to explain this evaluative judgment by deducing it, say, from the very concept or definition of music itself or by reference to this non-natural property called 'expressiveness' that the aria exemplifies so well. In both cases we would be denying the evaluative judgment the autonomy it should properly have. But this is not to endorse arbitrariness in ascription of evaluative terms. Not *any* piece of music can be properly called 'expressive', just like not *any* action can be labeled 'cruel'. We can correct the use of these terms against the appropriate paradigms, elucidate their meaning by pointing to examples, and we can be challenged to prove our understanding of evaluation by defending its application in a given situation.[151]

I suggest that we accept the above characterization of the nature of evaluation in general, and apply this reasoning to the case at hand. Our common reaction to amoralists as being insensitive, cruel, inconsiderate, treacherous and utterly selfish can be justified as *meaningful* by pointing to *examples* of insensitive, cruel or selfish actions on their part, and does not

---

[150] *Ibid*., p. 321.

[151] Davies and Jamieson (1987) in their response to Ross have argued that the autonomy of evaluative concepts need not be threatened even if we can give a successful analysis of these concepts in some other terms. The analogy they use is that of heat being analyzed in terms of kinetic energy, which, they suggest, does not yet require that our ordinary talk about heat should be replaced by talk about kinetic energy. But reference to 'ordinary talk' seems to be misguided in this case. Surely the concept of heat has lost it 'autonomy' in *scientific discourse* ever since its genuine nature has been discovered.

require the adoption of a particular metaphysical theory about the nature of human beings in general or the ontological status of moral properties.

## 4. What Does it Mean to Lack Morality?

There appear to be two senses in which a person might lack morality, both of which are closely related to each other. First, a person might accept a system of action-guiding principles which he (mistakenly) takes to be a system of *moral* (as opposed to non-moral) principles, but which, as the matter of fact, fails to satisfy the criteria necessary for being *a* morality. An ethical egoist, I shall argue, is the most familiar type of an amoralist who lacks morality in the sense of behaving in accordance with the system of non-moral principles. Secondly, one might have no delusions that a maxim that underlies one's action at any given time is of a type that is part of the greater system of morality, however unconventional this morality might be. He might even agree that ethical egoism or related views are *not* alternative moralities, and he might be well aware of the requirements of some existing system of morality (e.g., religious ethics). What will still be missing, however, is a willingness to prefer the moral values to the non-moral ones. I shall refer to this case as a case of moral indifference, and I shall argue that this sense of lacking morality is more fundamental as compared with the first one.

### 4.1 Adopting a Non-moral Action Guide

In this section I shall argue that any view that arbitrary selects either a single person (e.g., oneself) or a group of people (e.g., one's own ethnic group) as the only proper recipients of the non-moral good or evil fails to satisfy the conditions necessary for being a system of morality.

The basic claim of ethical egoism is that one morally ought to promote one's own interest, even if promoting one's interest may adversely affect the interests of others. It is sometimes argued that ethical egoism, when carefully formulated, can be taken as a possible *moral* position, and there is nothing incoherent in discussing ethical egoism alongside the *prima facie* altruistic theories of morality, e.g., utilitarianism or Kantian ethics. It is rarely that philosophers accept ethical egoism as the *best* moral theory, but many are prepared to acknowledge its legitimate status as a contestant in the debate between various normative theories. Thus Novell-Smith writes:

> If a man regularly decides that he ought (in the verdict-giving sense of 'ought') to do whatever brings him pleasures or profits, his dominant pro-attitude is toward his own pleasures or profits. Whether or not we choose to call selfishness a moral principle with him, depends on the criterion we are using for the phrase 'moral principle.' If he behaves selfishly without acknowledging his wickedness and without feeling remorse, we could say that selfishness was one of his moral principles.[152]

In what follows I shall present two arguments against the contention that ethical egoism can be viewed as a moral theory. But we should start with a clear formulation of the thesis of ethical egoism first.

Out of the all possible formulations only two appear plausible as moral theories. *Personal Ethical Egoism* is the claim that I [the person making the claim] ought to do what will most promote my own interest, regardless of what happens to anyone else. *Universal Ethical Egoism,* on the other hand, extends the scope of the prescribed selfish behavior to cover all rational agents. It is the view that everyone (morally) ought always to promote their own interest, and disregard the interest of anyone else (unless, of course, caring for the interests of others may increase the agent's own welfare in the long run).

---

[152] Novell-Smith, 1954, p. 310.

We can see right away that Personal Ethical Egoism fails to satisfy the minimal requirement of universalizability which is partially constitutive of the moral point of view. As a personal egoist, one is committed to the course of actions that would maximize one's interest, he *cannot* will, *qua* personal egoist, any type of action or approve of any rule that would deter his self-interest. Allowing *others* to satisfy their interest too, however, will necessarily conflict with satisfying one's own interest. Thus a personal ethical egoist cannot consistently will to universalize his basic prescription to allow others an unrestricted pursuit of their self-interest, without thereby turning into an altruist or a universal ethical egoist.

The claim that *everyone* ought to maximize their own interest is a variety of egoism that is seemingly consistent with the universalizability requirement. As such it is probably the most plausible candidate for being a genuine moral theory. And indeed, depending on the background justification for this thesis, it can be taken as a form of utilitarianism. Adam Smith's conception of the classic capitalist society, for instance, implies that in a free economy the pursuit of individual economic interests is beneficial for all *in the long run*, and, as long as the general prosperity of society is something to be prized, everyone ought to (morally) pursue their own interest in their everyday business transactions.[153] This twist, however, is far removed from the original spirit of ethical egoism.

I shall thus consider the thesis of universal ethical egoism as long as it remains an *egoistic* theory, and evaluate its claim to being a *moral* theory. Two arguments will be advanced to show that even the most plausible formulation of ethical egoism fails to satisfy the criteria necessary for being a moral theory.

---

[153] Smith, Adam, "The Wealth of Nations."

Kai Nielsen argues that the notion of universal ethical egoism is an incoherent one since it defies the main function of morality.[154] Following the classical tradition, Nielsen contends that the main function of morality (i.e., why we need morality at all) is to allow some way for adjudicating conflicts of interest. There might be several equally successful social arrangements or normative systems which would allow avoiding an open conflict and fairly and consistently resolve the issues of property, arrange personal relations, and the like. Each of these systems would correspond to a different morality. But it is *not* the case, according to Nielsen, that universal ethical egoism is just another alternative from the range of all possible systems of morality:

> We should now apply this consideration back to the problem of egoism. If, where the interests conflict, we are simply told that each person is always to act to satisfy his own interests, we have no way rationally to adjudicate conflicts of interest. We in effect tell people who are in conflict to go on conflicting. […] Then it is plain that we have no ethical theory at all, for so-called ethical egoism fails to do what most centrally an ethical system exists to do, namely to adjudicate in a fair and reasonable manner conflicts of interest.[155]

One can still attempt to save the theory by claiming that in a society where every member is a rational or enlightened ethical egoist there would be *no* conflicts of interest.[156] And thus adopting universal egoism is tantamount to resolving all possible conflicts once and for all. It would follow then that ethical egoism, far from being incoherent and destructive of societal existence, fulfills the main function of morality *par excellence.*

---

[154] Nielsen, K., 1989, p. 163ff.

[155] *Ibid*., p. 164.

[156] Ayn Rand, for instance, argues that "rational interests do not clash." ("The Virtue of Selfishness," 1965, p. 57).

In response, it is sufficient to point out the implausibility of this utopian scenario of universal harmony. If the claim about the perfect compatibility of rational interests is meant to be an *empirical* observation of the real state of affairs, it can be easily falsified by pointing to a number of actual conflicts where both sides of the conflict seem to be in their right to pursue their own conceptions of the good, and without any side lacking in rationality in any plausible sense of this word (e.g., the conflict of interests that occurs during labor negotiations). If, on the other hand, the above claim is presented as an *a priori* truth, where the notion of compatibility is built into the definition of 'rational interests', it is hardly informative, and simply begs the question against the objector.

But there is a second argument that questions the basic assumption that underlies all of the versions of ethical egoism. I shall adopt this argument with some modifications from James Rachels.[157] All possible formulations of egoism have this one feature in common – they *single out* one person for a special, preferential treatment (say, insofar as the distribution of non-moral good is concerned), and the basis for such preference is simply that one enjoys a very special relation with that person (namely, the relation of identity). Whether the theory permits that all others assume this preferential policy towards oneselves as well or not, makes no difference in this case. Whoever utters the principle of ethical egoism agrees to the preferential treatment of oneself simply on the basis of being oneself.

But one can show that this attitude goes against the first condition of taking a moral point of view, namely, the universalizability condition. One of the immediate implication of the first condition is that a difference in evaluative judgment cannot be justified by numeric difference between persons or situations. As James Rachels puts it, "We can justify treating

[157] Rachels, J. 2000, pp. 87ff.

people differently only if we can show that there is some factual difference between them that is relevant to justifying the difference in treatment."[158] Indeed, we can take it as one of the fundamental principles of morality. I have suggested above that the ability for abstract reasoning and *fragility* are the two relevant features that can occasionally justify the difference in treatment and in moral judgment of different persons. These, however, are (usually) not the features that an egoist refers to when he demands a special treatment for himself. Thus, we may conclude that an action-guide assumed by an egoist is not an alternative morality, but rather a non-moral life-policy.

We can generalize the above reasoning to exclude from the moral realm a number of related views. Racism and ethnocentrism, for instance, are the views that demand preferential treatment for some limited *group* of people, rather than a single person. But the explanatory basis for this discrimination (such as origin, complexion or gender) is for the most part irrelevant to morality. There appear to be no morally relevant factual difference between the people of different ethnic groups, races or sexes that would justify difference in treatment. Thus, all forms of discriminatory ideologies are not genuine moralities.[159]

4.2 Moral Indifference

In the previous sections I have characterized the constitutive features of the moral point of view which would delineate the set of all possible moralities from some of the non-moral perspectives and non-moral action-guides. I have used this substantive characterization

---

[158] *Ibid*., p. 89.

[159] Ethical egoism and other discriminatory ideologies are by no means the only possible non-moral worldviews. Anthony Skillen, for instance, defends a view that Marxist philosophy implies a form of *anti-moralism*: "Marx's scorn for 'morality' should not be interpreted in terms of rejection of the prevailing (bourgeois) *content* of moral norms, which would be consistent with espousing an alternative set of 'precepts' or 'principles.' Rather, Marx's position should be seen as a theoretical and practical opposition to the very *form* of 'morality' as he understood it: the form of universal, absolute laws binding on individuals as beings with a capacity to rise above and conquer their selfish, capricious 'inclination.'" (1981, p. 156). If Skillen's argument is acceptable then a consistent Marxist would also be an amoralist.

of morality to test the claim that ethical egoism (and related discriminatory views) constitutes an alternative moral code, and have argued that it fails this test. All this might be conceded by an ethical egoist. Accepting my argument might prompt him to slightly alter his terminology. If he is now convinced that there are definite rules which restrict the application of the term 'morality' only to action-guiding principles with certain features, which ethical egoism lacks, he will now avoid saying that everyone, *morally speaking*, ought to pursue their own interest, but will express it rather as a requirement of *prudence*. Even if the principle of ethical egoism turns out to be a non-moral guide, still, one might say, one will be better off by adopting *this* principle and disregarding the moral principles proper. Why should one care, that is, whether his life-style corresponds to some abstract definition of morality? The maxims that underlie my actions and the evaluative judgments that I occasionally make are non-moral - so what?

In order to better understand this possibility of moral indifference we might draw an analogy with religious attitude first. A person who does not accept the beliefs of a particular religion nor shares its values can nonetheless acquire (intellectual) knowledge of the factual criteria used in making evaluative judgments about actions. There is nothing, it appears, that prevents a non-believer from being able to identify correctly certain actions as sinful from the point of view of that religion, and it would be paradoxical to suggest that his (correct) judgment is somehow fundamentally different from a judgment with the same content but which comes from a believer. In a similar vein, Ronald Milo argues that we must distinguish between saying, from the moral point of view, "X is wrong," and saying, "X is wrong from the moral point of view."[160] If (as it was argued) the correct application of the moral

[160] Milo, 1984, p. 178. Milo himself attributes this distinction to Frankena.

predicates to actions (logically) depends on the recognition of the morally relevant facts (i.e., facts about the well-being of sentient creatures), there seems to be nothing contradictory in a scenario where someone reliably identifies the morally praiseworthy or blameworthy actions, and makes a number of genuinely moral judgments (i.e., he is willing to universalize these judgments in the relevant sense, and his reasons for a judgment deal with the facts about human welfare), without yet *adopting* the moral point of view himself. Thus, Milo writes:

> Even the so called 'amoralist' - i.e., the person who is himself indifferent to matters of moral right and wrong – can make genuinely evaluative judgments about what is right or wrong. Indeed, if the 'amoralist' is a very subtle thinker, he might be called upon by those who are morally concerned to advise them about the moral propriety or impropriety of such things as abortion or 'reverse discrimination.' And if the 'amoralist' understands the criteria that define the moral point of view and understands (intellectually) the feelings and attitudes that make others concerned about whether their behavior conforms to these criteria, then he may choose (perhaps for a fee) to advise them and hence engage in the process of moral evaluation. If he now employs his criteria to govern his judgments about what is right and wrong, he will be judging from the moral point of view.[161]

If our argument for descriptivism is cogent the above scenario should at least be seriously considered. There was nothing in the four conditions discussed that required the one making a moral judgment to give, in addition, the *unquestionable preference* to the judgments and prescriptions of morality as opposed to any other normative judgments and prescriptions. Hence, Milo continues, "there is no reason why a person could not judge it to be a morally bad thing to slaughter people on the lawn merely in order to create the aesthetically pleasing

[161] *Ibid*., p. 179.

contrast of red on green even though he himself prefers the creation of beauty to the avoidance of moral wrongdoing."[162]

Milo's strongest argument for the coherence of the above scenario is the appeal to the common intuitions. It seems to be an essential part of the ordinary understanding of an evil or amoral person that the problem in this case is primarily with his (perverted) system of values or priorities (motivational factors) rather than with his (mistaken) beliefs. Having the correct moral beliefs, including the beliefs that certain universal moral principles are true, is not yet sufficient for *caring* about morality. What still might be lacking is the overriding motivation to prefer (in cases of conflict) the course of action dictated by the system of morality to the course of action dictated by some other non-moral normative system or considerations of self-interest. Thus, pointing to indifference to (known) prescriptions and prohibitions of morality is the most natural way to describe the sense in which someone might lack morality.

Two objections to moral indifference thus described must be briefly considered. First, Milo's characterization of this kind of amoralist suggests that understanding morality involves grasping the truth of certain factual beliefs about the moral facts. It implies, in other words, the truth of *cognitivism*, and begs the question against the *non-cognitive* interpretations of morality, where moral attitude is identified with having a certain *attitude* toward the world. Thus, on the non-cognitivist interpretation, one cannot have a purely intellectual grasp of the moral beliefs and principles, since there is nothing that can be so grasped. To be sure, one can acquire the true beliefs *about* the attitudes of a morally concerned person, but not the true moral beliefs themselves.

The second charge deals with the alleged motivational efficacy of moral beliefs (or attitudes). Motivational *internalism* is a view that asserts a conceptual or logical connection

[162] *Ibid*., p. 180.

between accepting the truth of a moral principle and acquiring (at least some) motivation to act upon this principle. Since non-cognitivists identify a belief that, say, "X is wrong" with a con-attitude that the agent has, all non-cognivists will also be motivational internalists. On the other hand, there are philosophers[163] who accept the cognitivist interpretation of moral beliefs but insist on the necessary connection between the belief that "X is wrong" and a motivation to avoid doing X.

If internalism is true, it makes no sense to say that John sincerely believes that X is morally wrong but does not care in the least about avoiding X. An internalist would reinterpret the situation by insisting that the morally indifferent John does not *really* make a genuine judgment that X is wrong, or that whenever he utters 'I believe X is morally wrong' he must be using the phrase 'morally wrong' in inverted commas, indicating that his claim is actually about what *other* people believe.

I want to suggest that my conception of amoralism, even though it most naturally fits within the framework of cognitivism and motivational externalism, is nonetheless compatible with both non-cognitivism and internalism. Thus there should be no conceptual, metaethical reasons to reject my subsequent explanation of amoralism in the following chapters. I have indicated in Chapter One that my understanding of amoralism is more inclusive and less demanding than that of Ronald Milo. As such it can accommodate a variety of metaethical views on the true nature of moral beliefs and the motivational efficacy of those beliefs.

The essential feature of an amoralist, on my view, is the absence of a desire to be moral, i.e., to do what is right from the moral point of view. We may concede to a cognitivist and externalist that a person may possess the intellectual knowledge of moral principles and skillfully apply these principles to particular acts in situations that demand a moral judgment,

[163] See Hare, R. (1981); Smith, M. (1994) and Thomas Nagel, *The Possibility of Altruism*, (1970).

without himself being motivated to give precedence to moral precepts over precepts of other normative system or requirements of self-interest. If, however, as the defenders of both cognitivism *and* internalism suggest, it is logically impossible for one to have a belief that "X is morally wrong" without developing a strong con-attitude against doing X, we should perhaps reject Milo's scenario of the amoral 'moral councilor', but still retain our initial understanding of amoralism as the state of indifference to moral requirements. Lastly, the truth of the claim of non-cognitivists that moral beliefs are not genuine beliefs that allow of truth and falsity but are rather conative states of approval or disapproval of certain actions and principles would force us to rephrase a description of the actual cases of amoralism in a way that would avoid references to moral beliefs, but would not invalidate the central insight. An amoralist lacks motivation to prefer moral values to other values *whether or not* he can also be described as having correct (and genuine) moral beliefs.[164]

At the end of this section I want to suggest that being *indifferent* to the requirements of morality is the most basic way in which one can *lack* morality, the one that ultimately underlies the cases of adopting a non-moral action-guide as well. It will be argued in Chapters Four and Five that this kind of indifference is the result of a mistaken value judgment rather than an inherent emotive deficiency.

It seems that a person may adopt a non-moral action-guide (e.g., ethical egoism or racism) as his life policy for at least two reasons. First, being motivated to do what is morally right, he might be sincerely mistaken by taking these normative systems as systems of *morality*. Secondly, he might be aware of their amoral (and immoral) status, but nonetheless

[164] On my own view, both cognitivism and motivational externalism are the most plausible positions on moral beliefs, and fit best with the phenomenon of amoralism. But presenting an independent argument for each of these positions is beyond the scope of this work.

*prefer* the non-moral system to the system of morality[165] (e.g., he might believe that amoralism pays better in the long run). In the latter case what accounts for his amoralism (superficially) is the lack of (sufficient) motivation to be moral (or to adopt a moral action-guide), and thus the explanation of moral indifference will *also* be the explanation of an amoralist in the first sense (i.e., as a *systematic* amoralist). In the former case, we have an instance of moral *ignorance*, where what is lacking is not the desire to be moral, but the knowledge of *what* morality is all about. It is a case of ignorance of the most fundamental (non-derivative) moral principles, rather the ignorance of the particular facts that results in misapplication of those principles. However, such ignorance is highly unlikely on the individual level, and impossible on the universal level (i.e., it is impossible that *everyone* is mistaken about what constitutes a moral point of view and which feature of the world are morally relevant). Thus, the possibility of a *conscientious amoralist* can be ignored in this discussion. Those rare individuals that are sincerely convinced that ethical egoism or anti-Semitism are the genuine moralities and rigorously follow all the requirements of these action-guiding systems are persons in need of moral instruction, rather than moral blame. Moreover, the explanation of *their* amoralism is on the surface (namely, a mistake in reasoning or ignorance) and there is no need for a deeper theoretical investigation.[166]

[165] To be sure, this need not be a momentous decision that determines the rest of one's life once and for all. Usually, the decision will need to be re-made at each occasion of conflict between morality and, say, self-interest.

[166] A much more common case is a case of immoral behavior due to mistake about the proper *application* of the moral principles, or the failure to realize that a particular act falls under the category denoted by the universal rule in the major premise. But those are *not* cases of amoralism.

# CHAPTER III

## THE FACES OF AMORALISM

> On that occasion, sitting at tea and chatting with the crew, for the first time in my life I clearly formulated the following for myself: I have neither the feeling nor the knowledge of good and evil, and not only have I lost the sense of good and evil, but good and evil really do not exist (and this pleased me) and are but a prejudice.
>
> Dostoevsky, *The Possessed* (Stavrogin)

The immediate purpose of this chapter is to present the somewhat abstract philosophical notion of the amoralist in its concrete real-life manifestations. This will further show that amoralism is not a theoretical phenomenon constrained to a merely possible world, but is rather a conspicuous part of everyday reality. The more strategic goal is to gather factual material necessary for the interpretative and explanatory task of the following chapters. Before we can meaningfully approach the explanatory stage of this project it is prerequisite that we have a reasonably complete description of an amoralist's typical behavior, beliefs, attitudes and dispositions. This chapter, then, secures a necessary factual foundation for the subsequent theoretical part.

### 1. The Extreme Amoralist

During the initial introduction of the amoralists in Chapter One, I suggested to make a distinction within this general group of individuals with the absent moral sense between the extreme and selective amoralist. This categorization will be explained and justified below by pointing to the specific cases of amoralism, which exemplify the features typical of each

kind. At the same time, it is understood that the usefulness of these terminological constructs should be assessed using the pragmatic criteria – their usage is defensible in so far as they simplify and facilitate the overall discussion of the phenomenon of amoralism in the context of this work's overall goals. As in cases of other terminological distinctions, no presumption is made that these categories capture the real divisions in the natural world, and it should be acknowledged that a number of alternative (and perhaps equally acceptable) conceptualizations is possible.

1.1 Who are the Psychopaths?

In the *Enquiry Concerning the Principles of Morals* David Hume claims that there has never existed a person who was totally indifferent to moral reasons:

> If we consider the principles of the human make, such as they appear to daily experience and observation, we must, *a priori*, conclude it impossible for such a creature as man to be totally indifferent to the well or ill-being of his fellow-creatures, and not readily, of himself, to pronounce, where nothing gives him any particular byass, that what promotes their happiness is good, what tends to their misery is evil, without any farther regard or consideration.[167]

Indeed, Hume believed that the fact that one "wears a human heart," i.e., belongs to the human race, would ensure the presence of a moral sentiment, however weak or suppressed, and it may strike one as being very un-Humean that he takes this statement as an *a priori* truth. However, having perhaps an advantage over Hume in that we have more than two hundred additional years of human history to rely upon, we may suggest with good reason that Hume's estimation of the innate human beneficence was perhaps overly optimistic. Not incidentally, Erich Fromm defines man as "the only primate that kills and tortures members

[167] Hume, D. (1777/1966), p. 264.

of his own species without any reason, either biological or economic, and feels satisfaction in doing so."[168] The following discussion is meant to illustrate this point to some extent.

It seems opportune to begin this discussion with the extreme cases of amoralism first. Whenever a concept is of such a kind as involving some vagueness about the extent of its proper application, dealing with the unambiguous instances first may help identifying the essential features that would apply to all members of the class, and the features that are perhaps only contingent and accidental. Being amoral is clearly a matter of degree and the examination of the most conspicuous cases of amoralism will set a proper background for the later discussion of the less pervasive, albeit no less shocking, manifestations of this condition.

I have suggested earlier taking a psychopath[169] as a paradigm example of an extreme amoralist. A brief explanation and justification of such a choice is, however, in order. One objection that can be raised at this point is that a psychopath is a character that properly belongs to the domain of psychiatry or clinical psychology, but hardly to a work in philosophy. Psychopathy is apparently a medical condition, that cannot be competently discussed, let alone explained, by someone lacking a formal education in the appropriate filed of medicine. A philosopher, on this line of reasoning, ought to confine his investigations to the fictional characters of the philosophical treatises and dialogues, and should therefore leave the 'real-world' cases to the empirical sciences, that are better equipped to deal with them.

---

[168] Fromm, E., 1977, p. 26. Cf. also Hannah Arendt's discussion of 'radical evil' in her "The Origins of Totalitarianism" (1951).

[169] The terms 'sociopath', 'psychopath' and 'antisocial personality' are used interchangeably here.

That a philosopher can benefit from the work of the empirical scientists, without being an empirical scientist himself, and usefully employ their results toward the solution of some genuine philosophical problem will hardly be denied. What usually raises a concern is the extent and quality of the contribution that a work in philosophy can deliver on the questions that have long been considered the rightful property of the 'hard' sciences. Various well-known historical precedents are often cited to show the futility and hidden danger of such ambitious pretenses on the part of philosophy.

There is much that can be said against such an unduly narrow view of the proper domain of the philosophical interest, and dogmatic insistence on the strict compartmentalization and autonomy of the sciences. But it will be enough here to point out that the above objection to the use of psychopathic personality as one of the main characters in a philosophical work is largely based on a common misconception of the phenomenon of psychopathy itself. Contrary to the popular opinion, a psychopath is neither a mentally ill person, nor a person suffering from any known defect of the intellect, psychosis, insanity or brain damage.[170] There is almost a uniform opinion that concurs with this claim starting from the earliest investigators of this condition in the beginning of the 19th century (Pinel, 1801; Rush, 1811; Conolly, 1830; Prichard, 1835) to the latest authoritative studies of the psychopathic personality by Hervey Cleckley (1982), Benjamin Wolman (1987; 1999), Robert Hare (1993), and Martha Stout (2005).

Psychopaths firmly belong to the category of philosophical curiosities, and the references to serial killers, mass murderers and other hardened criminals are often used to provide vivid illustrations for the discussion of the dangers of moral skepticism or, perhaps,

[170] Except for rare cases (see Ch. 5).

the consequences of a materialistic worldview. Nonetheless, psychopathy remains among one of the most puzzling social phenomena. It is well known (e.g., from popular literature and films) that psychopaths can be quite charming and often exhibit a level of intelligence that is significantly above average;[171] they are able to achieve a high social status and to function successfully in various social arrangements for extended periods of time. As Stout observes, this peculiar "disfigurement of character" that she calls 'sociopathy' is unique in that

> …all of the other psychiatric diagnoses involve some amount of personal distress or misery for the individuals who suffer from them. Sociopathy stands alone as a 'disease' that causes no *dis-ease* for the person who has it, no subjective discomfort. Sociopaths are often quite satisfied with their lives.[172]

Wolman argues that "sociopathy seems to be more a social than a medical problem."[173] And Cleckley reports that psychopaths, when brought to the psychiatric facilities, are routinely discharged upon closer examination, since most of them are absolutely healthy as far as the accepted standards of mental health are concerned. As a result, he continues, there exists not a small confusion among the medical practitioners, social workers and the law enforcement officials as to whether a psychopath belongs to a psychoanalyst's chair, psychiatric ward, priest's office or a prison cell.[174]

This shows that, at present at least, no single science or field of study can claim exclusive rights to a psychopath. As a complex social phenomenon, the psychopathy defies any confinement within the borders of one discipline, and only the collective, mutually

---

[171] Contrary to what the term 'psychopath' suggests etymologically, viz., '*sick mind*'.

[172] Stout, 2005, p. 12.

[173] Wolman, 1987, p. 22.

[174] Cleckley, 1982.

enriching work of many specialists from different fields would shed light on the causes of psychopathy, and can perhaps point in the direction of a solution to this problem.

While the psychopath is a rewarding and legitimate subject of study for a variety of disciplines, it is hardly surprising that the ways in which the investigation of these individuals is conducted, and the methods used, will differ significantly from field to field. The approach that is characteristic of a clinical psychiatric study of psychopaths, will be very different from the one assumed by an anthropologist, and even more distinct from the philosopher's take on this issue. Which one method, if any, will prove to be the most fruitful and successful in the long run remains yet to be seen.

This being said, we may proceed by introducing the amoralist in his extreme manifestation. As it was said above, I have chosen a psychopath as a paradigm example of an extreme amoralist. The two notions are not fully coextensive, as some autistic and narcissistic individuals may exhibit many of the features of the extreme amoralist as we have defined him, without yet being psychopaths[175]; but it is sufficient for our purposes to describe what is probably the most paradigmatic and the most intriguing type of amoralist.

Given the complexity and versatility of the condition of a psychopath as this type is described in psychiatric literature, it would be rather unhelpful trying to capture it by a single definition. The great variation that exists between the concrete cases of psychopathy suggest of adopting a 'family resemblance' model, where the members of this 'species' are loosely connected with each other by several overlapping and crisscrossing character traits, without (necessarily) there being any single characteristic that all members of this class share. This does not mean that we cannot elucidate this condition by citing the several most typical

[175] Cf. Sacks, 1995 (last essay); Grant & Boucher, 2005.

features of a psychopath. Martha Stout, a psychiatrist and sociologist, argues that "sociopaths, by various names, have existed in all kinds of societies, worldwide and throughout history."[176] And an anthropologist Jane Murphy reports that during her extended field work among the Eskimos of Alaska, and the Yorubas tribe in Nigeria she discovered that these two very different communities have a concept that closely parallels our concept 'psychopath' – someone who consistently violates the accepted norms of society and lacks the necessary emotive apparatus to feel genuine guilt or remorse over his deeds. The Eskimos have a word *kunlangeta*, which refers to a person "whose mind knows what to do but he does not do it."[177] A *kunlangeta* person consistently lies, cheats, steals and refuses to go hunting, and when other men are out of village, he takes sexual advantage of many women. As Murphy observes, the awareness of the rules is not in question here, and the violations happen despite numerous reprimands from the elders and are due to reasons other than ignorance or negligence. The Yorubas of Nigeria, too, have an abstract word with similar meaning, *arankan*, which refers to a person "who always goes his own way regardless of others, who is uncooperative, full of malice, and bullheaded."[178] A pattern of behavior characteristic of a *kunlangeta* in Alaska and an *arankan* in Nigeria is similar enough to the various descriptions of the extreme amoralists in our own culture to elicit a strong sense of cross-cultural reality of this phenomenon.

In a helpful discussion of the psychopathic or antisocial personality Richard Jenkins offers this brief characterization: "The psychopathic personality is not so much a disorder as

---

[176] Stout, M. 2005, p. 136.

[177] Murphy, Jane. 1976, p. 1026.

[178] *Ibid*.

it is a defect in development. In one sense we are all born psychopaths. Most of us outgrow this stage of overwhelming egocentricity: some become arrested there."[179] He further adds that "the psychopath is simply a basically asocial or antisocial individual who has never achieved the developed nature of *homo domesticus*."[180] The key character trait that is identified here by Jenkins as indicative of a psychopath is "overwhelming egocentricity", and the antisocial or even criminal behavior that is typical of such individuals is likely to be derivative from this main feature. Robert Hare (1993) and Martha Stout (2005) stress the absence of *conscience* as a distinctive feature of psychopaths, where conscience, "the most evolved of all humanizing functions,"[181] is understood in part as the ability to transcend one's egotistic desires and preferences, to consider the interests of others in making a decision, and experience guilt and remorse when others are adversely affected by one's actions. In the same vein, David Lykken observes that "the psychopath is characterized by a lack of the restraining effect of conscience and of empathic concern for other people."[182] These authors seem to agree that pathologic egocentricity and, as a result, total lack of empathy for others in distress sets the psychopath distinctly apart from the ordinary people, even from people with little inclination toward altruistic acts or the ones of reserved and apathetic disposition. An extreme amoralist, on most accounts, markedly stands out from the usual cases of the indifferent, wayward or criminal individuals. Dr. Hervey Cleckley, a psychiatrist who worked for many years with the individuals diagnosed with psychopathic personality reports on a case of a young man, named Gregory. Cleckley carefully records the many instances of

[179] Jenkins, 1960, p. 324.

[180] *Ibid*., p. 331.

[181] Stout, 2005, p. 10.

[182] Lykken, 1995, p. 115.

Gregory's antisocial acts (thefts, violence, promiscuity, lies, impersonations, etc.), but finds the most telling feature of his personality that defines him as a psychopath (an amoralist), and clearly sets him apart from the ordinary criminals (immoralists, as we might say), in the following:

> Gregory's utter lack of understanding why it might be considered just and appropriate for him to go to the state prison for committing crimes similar to those committed by others confined there well illustrates a major point that I think often distinguishes the psychopath from other people who carry out criminal acts. Other criminals do not, of course, want to go to prison, and often protest against it. But they do not seem to have this strange conviction that they are, or should be, somehow exempt from prisons that were made to control people who commit the very crimes of which they themselves have been convicted.[183]

What is most interesting in this observation, given by a psychiatrist rather than a philosopher, is a reference to a lack of the feature that we have identified as being necessary for acting within the moral domain, viz., willingness to universalize one's normative judgments. I have argued that on a minimal interpretation of this condition, one ought to apply similar evaluative pronouncements in similar situations or to similar persons, unless there is an essential difference between them. In Cleckley's example, Gregory's attitude is conspicuous for his unwillingness to apply the same criteria of evaluation to himself as would be appropriate (even in his own judgment) to others committing similar offences. Moreover, as Cleckley observes, this unwillingness on Gregory's part is very different from the understandable unwillingness of a criminal to face the penalties for his actions. A criminal would ordinarily agree that anyone committing *X* should get a certain penalty, but will on occasion argue that he did not in fact commit *X*, or plead some special extenuating

[183] Cleckley, H., 1982, p. 88.

circumstances that should spare him the punishment. The defense tactics of Gregory is not of this kind, according to Cleckley. Gregory is quite convinced that he should not be punished for doing *X* simply because it is *he* who did it, not the other person.

This fundamental difference in *thinking* is what apparently prompts Jenkins to go as far as to exclude the psychopaths completely from our own 'species' (conceived under the social aspect) – *homo domesticus*.[184] In what follows I shall illustrate these characteristic features of an extreme amoralist by drawing on two different sources. First, I shall present a summary of a period of life of two real individuals, whose grossly maladjusted social lives and peculiar thinking patterns betray the typical symptoms of a psychopath. These case-studies will be taken from the rich factual material collected by Dr. Hervey Cleckley over the many years of his professional career as a clinical psychiatrist and a professor at the University of Georgia. Secondly, I shall consider a character from a work of fiction. Nikolay Stavrogin from Dostoevsky's novel "The Possessed" presents arguably one of the most salient manifestation of the extreme amoralist personality that can be found in classic literature. In both cases the emphasis will be on the descriptive presentation of behavior (one actual, and the other fictional), and interpretation of these facts will be reserved for the following chapters.

1.2 The Two Case-Studies: Anna and Joe

A word needs to be said why mainly *non-violent* psychopaths were chosen for this presentation, whereas the word 'psychopath' is strongly synonymous in ordinary speech with serial killers, individuals committing extremely atrocious and senseless crimes and the kind

[184] This attitude is not uncommon. Jeffrie Murphy (1972) argues that "from the moral point of view, it is very implausible to regard [psychopaths] as *persons* at all… A psychopath is more profitably pictured as an *animal*." (pp. 293-4). Similarly, Harry Frankfurt (1971) would deny personhood to psychopaths on the ground that they (apparently) lack the second order desires (and thus belong to the class of *wantons*).

of people described by Aristotle under the name 'brutish characters'. Stephen Giannangelo in his study of serial murderers and Stanton Samehow in his investigation of a criminal mind agree that many violent criminals do indeed exhibit the characteristics of a classic psychopath. Thus Giannangelo stresses that "labels such as antisocial, borderline, narcissistic, and psychopath do apply [to serial killers],"[185] and Samehow goes as far as identify the persistent criminal with the psychopath.[186] But there are sound statistical reasons to think that the violent type of the psychopath is not representative of the general population of people with antisocial personality disorder, but rather constitutes an exception to the rule. While most psychopaths are quite susceptible to petty crimes such as theft and forgery, very few will resort to open and unusual violence and deteriorate into the obsessive and perverse murderers that are so familiar from popular films, books, and reports of the actual trials. But just this consideration alone should settle the issue: Stout estimates that as many as 4% of the population can be classified as sociopaths,[187] and Cleckley similarly (but more cautiously) states that "it does not seem an exaggeration to estimate the number of people seriously disabled by the disorder now listed under the term *antisocial personality* as *greater than* the number disabled by any recognized psychosis except schizophrenia."[188] It is obvious, however, that the percentage of serial killers and other criminals of pathological brutality is significantly less. The fact that the latter characters get much more publicity in the media may account for the widespread (but distorted) stereotype of the typical image of a psychopath.

---

[185] Giannangelo, 1996, p. 19.

[186] Samehow, 1984, p. 181.

[187] Stout, 2005, p. 6.

[188] Cleckley, 1982, p. 14.

From the many case studies of real-life psychopaths collected by Cleckley, the story of Anna strikes the reader as especially remarkable.[189] She was a woman of good looks, exceptional intelligence and substantial education whose favorable external appearance made it hard to believe some of the well-authenticated facts about her tumultuous life. In a brief layout, some of the most indicative incidents from her life story can be presented as follows. She was born to a relatively prosperous family in Georgia, and there were no indications that her parents were anything less than respectable and decent members of the community, who brought her up in the best traditions. Already as a high-school student, however, it was discovered that she was in the center of a peculiar boy's club, all the members of which were united by their shared sexual experiences with her. The ever expanding membership of this club finally led to its scandalous discovery, and after understandable commotion, Anna was sent to a boarding school in a distant state by her parents.

There were excellent reports on Anna's behavior and progress in studies for several months, with several teachers commenting on her high intellectual abilities. However, the eventual accumulation of various rule violations (including numerous petty thieveries) led to her expulsion. She has subsequently attended half a dozen schools, remaining nowhere for more than half a year. The typical kinds of incidents that led to her expulsions from schools include inappropriate and offensive pranks, extreme rudeness to teachers, forging grade reports, occasional 'borrowing' of cars belonging to teachers or other school employees without asking, pick-pocketing and overactive sexual behavior. On one occasion she outraged several girls in her dormitory, with whom she had been friendly, by coming to the room in their absence and urinating on their evening dresses which they highly valued. This incident was apparently not provoked by any animosity or recent argument between them,

[189] *Ibid*., pp. 48-61.

and the motivation behind it remains unclear. What is significant though is her subsequent reaction to it: "On being confronted with plain proof from several sources [that she did it], she was able to smile off the affair and dismiss it as a whimsical prank."[190] In a similar manner, the numerous stealing incidents, from pocket change to cars, could not be explained by any real need on Anna's part. Most of them are best described as actions done on some trivial and fleeting impulse that she was unable or unwilling to control.

Anna's later life was characterized by the succession of marriages and divorces, too numerous to mention. As Cleckley observes, "she had fallen into the habit of marrying on an impulse apparently as trivial as what might lead another woman to buy a new dress."[191] Her husbands were adventurous taxi drivers, bar companions, random people she would meet at work, and other equally questionable companions. At the same time, her married status did not in any way constrain her sexual activity with other men. While being married to a man in England, "to him and almost simultaneously to his wayward brother Anna is said to have given gonorrhea, as the first year of marriage drew to a close, she inadvertently picked up from a sycophantic interior decorator, a person, more active homosexually than otherwise."[192]

Cleckley emphasizes that it was not her consistently immoral behavior that was the most astonishing about her, but rather her indifferent attitude toward it: "On being detected in activities that would produce fear, shame, or consternation in others, this patient often showed simple insouciance."[193] This lack of emotional involvement on her part (her ability to

---

[190] *Ibid*., p. 56.

[191] *Ibid*., p. 58.

[192] *Ibid*., p. 59.

[193] *Ibid*.

escape "the subjective consequences of her experiences") and total disregard for (or unawareness of) the feelings of others are features characteristic of a psychopathic personality. There was no evidence that any purely cognitive deficiency was involved, or that her understanding of the norms and requirements of common morality was perverse. Interestingly enough, during her long and versatile career she, among other things, taught a Sunday school class, and according to observers, "her teachings were ethically admirable and she gave a strong impression of sincerity."[194]

The basic details of the early life of another subject, Joe, are in many respects similar to the previous case. Joe's father was a prominent man in Alabama of puritanical ideals, who gave his son the very best education. The school teachers characterized him as "unusually bright," and he eventually graduated with distinctions from law school, being named valedictorian of his class. Already during his studies he began "to behave outlandishly when he drank", often ignoring his serious duties on which his own success in school vitally depended. After law school, periods of heavy drinking and irresponsible behavior were alternating regularly with times of relative stability and sound attitude. He became interested in running for city council, and after the most active and inventive campaigning, won the elections. But right after that he lost any interest in the position (which took him considerable effort to obtain), neglecting even the basic duties, and it eventually did cost him the job.

While still being a city council, he often acted in extraordinary and distressing way, that truly puzzled those who saw him. On one occasion, being a guest at a formal dance, Joe startled all those present by coming toward an attractive young lady, jerking her dress up over her head, and "divesting her of her undergarments, and, despite her struggles and

---

[194] *Ibid*., p. 60.

screams, commenced on his first steps toward an attempt at sexual intercourse in these inauspicious surroundings."[195] Only the considerable number of eye-witnesses of good reputation made it possible for many who knew Joe, and had the best possible impression of him, to believe such an extraordinary story.

Joe did marry, but his marriage was a failure from the start. Long and unexplained absences from home, neglect, multiple and thinly disguised affairs with other women and the total lack of concern for his wife's feelings were things typical of Joe's brief marital experience. He was able to secure a position of a judge of a local court, largely through his father's efforts, but he soon lost his job being unable or unwilling to attend even to the barest minimum of his duties. Joe was equally irresponsible with his money, as in all aspects of his life, "wasting all the saving and his ample income on the most frivolous adventures." However, as Cleckley notices, "there seemed little incentive or definite purpose in the actions through which he destroyed his opportunities and squandered his and his parent's resources."[196] What surprised the observer most, was that there seemed to be no strong recognizable temptation, and no clear life goal, good or bad, that would provide motivating reason for abandoning his job, his wife, his children and everything else he apparently valued.

Naturally, doubts about his mental condition did arise, and Joe was sent to psychiatric hospitals on several occasions. But no sign of psychosis or psychoneurosis could be found. He was usually classed as a case of psychopathic personality and promptly released from the hospital. He fell into the hands of police more than once for petty fraud and similar

[195] *Ibid*., p. 73.

[196] *Ibid*., p. 74.

misdemeanors, but "always disarming and impressive, he cleverly talked himself out of the usual consequences."[197] Wandering from city to city and following an odd impulse, he walked into a tent of a celebrated evangelist, impressed the preacher and all around with his apparently sincere conversion and enthusiasm, and traveled with the group for several months, being "active and successful in bringing in the penitent and getting them out on the sawdust trail."[198] His new career as an evangelist ended abruptly and in the usual manner: after meeting an attractive woman, he left the group, stayed with her in a hotel for some time, and "abandoned her as casually as he had begun the relationship."

Joe's subsequent life is a long record of impressive 'fresh starts' in different parts of the country, and equally impressive failings. He worked for a short time as a promising journalist in New York, but soon lost interest and incentive for work, associated himself with a number of delinquent groups, frequented bars, ceased all efforts to write and essentially wasted another good opportunity to establish himself in a socially acceptable position. Upon returning home, and behaving in his customary way, he was once again referred to a psychiatric hospital, where he gave impression of an absolutely sane and reasonable man, who regrets his previous failings. He was eloquent in outlining his plans for a new and radically changed life and promised with all signs of sincerity to abide by his new commitments. A parole was granted soon after, and in less than two weeks Joe returned to his usual pattern of unprovoked antisocial acts, drunkenness and self-defeating behavior.

Dr. Cleckley, who observed Joe for a period of time in a psychiatric hospital, reports his impression about Joe's inner state in the following words:

---

[197] *Ibid*.

[198] *Ibid*., p. 76.

> This man's apparent insight, freedom from evasiveness, and willingness to admit himself responsible for his misfortunes are inconsistent with his history, which is typically that of a psychopath. […] In time, his insight comes to seem but a mimicry of insight. He uses words that one who understood would use, but they do not have a corresponding meaning. He speaks with every evidence of conviction and sincerity, but when one studies him over a period of time, it becomes apparent finally that he is merely going through the motions, that he is not actually living the feelings he describes so well.[199]

Cleckley attributes this conspicuous incongruity that exists between Joe's behavior and his verbal professions of guilt, regret and sincere plans to reform to the fact that he "does not experience real and serious emotions." Indeed, it appeared that, although otherwise always inclined to lying, in many cases there was no *intentional* deception or acting involved on those occasions when Joe claimed to feel fully responsible for the situation and claimed to experience remorse or guilt over his actions. Rather, the situation was similar to the case of a color-blind man (without knowledge of his limitation), "who after the investigation swears conscientiously that the horizon is gray, though it actually blazes with all colors of the sunset."[200] It appeared as if an important dimension of human experience (non-cognitive in nature) was lacking from Joe's life, and there was no way to fill this gap by any other means.

These two brief summaries of some events in the lives of Anna and Joe show people who are involved in a clearly self-defeating pattern of behavior. Their goals, as revealed in conversation, are usually clearly defined and socially acceptable (e.g., finishing college, getting a job, retaining the position), and there is no lack of understanding about the means necessary to achieve the goal, as well as about the kind of actions that would prove detrimental to this effect. Nor is there lack of technical competence (general intelligence) that

---

[199] *Ibid*., p. 70.

[200] *Ibid*., p. 71.

may in other cases contribute to the failure. Clearly, the explanation of this "puzzling and spectacular failure to experience life normally and to carry on a career acceptable to society"[201] must be sought elsewhere, perhaps on the deeper emotional levels of one's personality or else, in that aspect of human rationality that is responsible for choosing ultimate values and establishing the stable hierarchy between them.

1.3 The Amoralist in Literature: Nikolay Stavrogin

It is often argued that a literary character, a creation of a great writer, may in some sense be more real than the actual flesh and blood person. This common and, in my view, justified impression is due to at least two factors. First, the writer is rarely working with purely fictional or imaginary material, but is often describing with great skill the inner characteristics of some real people that he knew and places them in situations and conflicts similar to the ones that he himself was involved in. Secondly, the literary character with all of his private thoughts, beliefs, intentions, and inner struggles is publicly accessible and can be easily observed in the most intimate details of his life, which, of course, is not always the case even with the patients who are under constant doctor's supervision, or the famous people under unceasing attention of the media. For this reason, looking in addition at the life of a fictional character, whose actions and beliefs betray the typical features of an extreme amoralist, would enrich this discussion, and provide further valuable material for the interpretive and explanatory efforts of the next chapter.

---

[201] *Ibid*., p. 252.

Nikolay Stavrogin[202] from Dostoevsky's novel "The Possessed" strikes the reader as one of the most unusual and intriguing characters. The only son of a general and a wealthy landowner he received a fine education in the capital city of St. Petersburg and, by his mother's wish, was preparing for a military career himself. Upon finishing school, he received a commission "in one of the most brilliant regiments of the Horse Guards."[203] To the great delight of his mother, Varvara Petrovna, Stavrogin, while staying in St. Petersburg, "renewed acquaintances which she [Varvara Petrovna] had hardly dared to dream of, and was welcomed everywhere with pleasure."[204] On all accounts, the young man was about to establish himself in higher society and secure a career that would provide respect and financial independence for the rest of his life.

But as is typical of the many actual cases of antisocial personalities, Stavrogin suddenly lost all interest in his original life-plans and resorted to the kind of conduct that cannot be easily explained by reference to the ordinary motivational factors:

> But very soon rather strange rumors reached Varvara Petrovna. The young man had suddenly taken to riotous living with a sort of frenzy. Not that he gambled or drank too much; there was only talk of savage recklessness, of running over people in the street with his horses, of brutal conduct to a lady of good society with whom he had liaison and whom he afterwards publicly insulted. There was a callous

[202] I am aware that a number of literary characters have been interpreted as embodying the most characteristic features of a psychopath. McGinn (1993) takes Yago from Shakespeare's *Othello* as a paradigm example of the extreme amoralist; Cleckley (1982) argues that Karamazov the father from *Brothers Karamazov* is the best example, Brink (1986) suggests Dicken's Uriah Heep, and Anthony Duff (1977) interprets the character of Meursault from Camus' *L'Etranger* in terms of antisocial disorder. Jeffrie Murphy (1972) is ready to assign "a strong psychopathic tendency" even to some apparently docile characters such as Ivan Ilych from Tolstoy's *The Death of Ivan Ilych*, whose life was devoid of any care and genuine concern for persons (cf. to the same effect F. M. Kamm's claim that Ivan Ilych belived that "it was alright to act on a maxim toward others that he would not be willing to universalize, including to have applied to himself" in "Rescuing Ivan Ilych." *Ethics*, 113, 2003, p. 204).

[203] Dostoevsky, 1936, p. 38.

[204] *Ibid*.

> nastiness about this affair. It was added, too, that he had developed into a regular bully, insulting people for the mere pleasure of insulting them.[205]

Stavrogin was involved in two duels, which he himself initiated, killed one of his adversaries, and severely maimed the other. In addition, it was said that he was living in strange company, associating himself with people far below his status, and apparently finding pleasure in that life-style. During that period of his life in the slums, "the time of debauchery" as he would refer to it in one letter, a lame and a mentally deranged girl, Marya Lebyadkin, who was staying in the poor quarters of the city with her drunkard brother, got infatuated with a well-dressed and well-mannered gentleman, as Stavrogin certainly was at that time. She saw Stavrogin frequently in his lodgings where she worked as a servant. Without giving much thought to it, and out of some strange obsession to violate all the norms of propriety and even common sense itself, Stavrogin secretly married the girl. As he would later explain his motivation, "the thought of Stavrogin's marriage to a creature like that, the lowest of the low, tickled my nerves; it would be impossible to imagine anything more monstrous."[206] We may note here that the property of 'being monstrous' was the reason for *doing* the action for Stavrogin. Shortly afterwards he left her alone with her abusive brother and in even more disturbed mental condition than before.

This string of antisocial and poorly motivated actions was apparently more than a temporary 'riotous' phase in Stavrogin's life in Petersburg. A number of equally incomprehensible and unjustifiable events occurred shortly after his return to his mother's estate, in a quiet provincial city. Only three will be mentioned here, by means of illustration. Having immediately impressed the local upper society with his exquisite manners and

---

[205] *Ibid*., pp. 38-9.

[206] *Ibid*., p. 713.

unfailing propriety with which he "performed all the duties demanded by the provincial etiquette", he very soon startled the public by a number of senseless and outrageous pranks, if one may call them such. One day being at the club, he went up to Mr. Gaganov, a respectable elderly gentleman, and "took him unexpectedly and firmly with two fingers by the nose, and succeeded in leading him two or three steps across the room."[207] The author specifically mentions that Stavrogin could have no grudge against Gaganov that could perhaps explain (if not justify) such behavior.

The second characteristic incident happened at a formal birthday party at Mr. Liputin's house. Stavrogin, being one of the honorable quests, took part in a dance with Liputin's young and attractive wife. The ordinary dance, however, had the most unusual ending:

> Stavrogin led out Madame Liputin – a very pretty little woman who was dreadfully shy of him – took two turns round the room with her, sat down beside her, drew her into conversation and made her laugh. Noticing at last how pretty she was when she laughed, he suddenly, before all the company, seized her round the waist and kissed her on the lips two or three times with great relish. The poor frightened lady fainted.[208]

Finally, Stavrogin's growing reputation as a man 'out of his mind' was solidified by his memorable encounter with Ivan Ossipovitch, the governor of the province and his distant relative on the mother's side. During the explanation with the governor at his office, Stavrogin acted as if he wanted to tell him something secretly to his ear:

> Poor Ivan Ossipovitch hurriedly and trustfully inclined his ear; he was exceedingly curious. And then something utterly incredible took place. The old man suddenly felt that, instead of telling him some interesting secret, Nikolay

---

[207] *Ibid*. p. 42.

[208] *Ibid*., pp. 46-7.

> has seized the upper part of his ear between his teeth and was nipping it rather hard.[209]

The governor would eventually need the help of his assistants to free himself from such an embarrassing position.

The several facts from the early life of Stavrogin in the capital, as well as these three curious incidents in his hometown, already show a personality unable to pursue any long-term life goals, with a peculiar and unstable system of values, as well as with clearly abnormal motivational schemes that underlie his behavior. Stavrogin's total disregard for moral and conventional rules of the society he lives in, his utter lack of emotive involvement in whatever he does, good or bad, apparent absence of inner constraints of conscience, his indifference toward the opinion of others together with a charming appearance and the obviously normal intellectual powers allow us to classify him (tentatively) as a case of psychopathic personality, and, consequently, as an extreme amoralist. Indeed, his story reads very much like one of the case studies collected by Dr. Cleckley, Robert Hare or Martha Stout. This first impression is further confirmed by what we learn (thanks to the author) about Stavrogin's inner world from his own "Confession", and especially about Stavrogin's crime of child rape.

Several researchers of the psychopathic personality have maintained that terrible *boredom* is one of the fundamental drives behind many of the antisocial activities that a psychopath is involved in. This constant feeling of boredom gives rise to the nagging need, indeed, an obsession, to search for ever new and more intense stimulations, many of them of a sensual kind. Wolman writes that "the 'pleasure principle,' that is, the principle of

---

[209] *Ibid*., p. 48.

immediate gratification of needs, is the main motive in the life of sociopaths."[210] A sociopath is never satisfied with his current state; he is always on the mission of searching for yet unknown pleasures and experiences. Once he reaches them, however, he gets bored quickly, and the same search continues again.

Stavrogin is no different in this respect from the diagnosed sociopaths. "I was so utterly bored that I could have hanged myself,"[211] he says during his confession at Tihon's; and again: "About that time I wished to kill myself because of the disease of indifference."[212] This terrible boredom, the 'disease of indifference' as he defines it, was apparently what moved Stavrogin to take sexual advantage of Matryosha, the only daughter of his landlords, a girl not yet twelve years old. This new experience, we can surmise, was his desperate attempt to escape (if only for a moment) the utter dullness and emptiness of his amoral existence.

After committing the crime of child rape, and, moreover, after *not preventing* Matryosha from hanging herself from shame and despair,[213] Stavrogin, on his own account, crossed completely to the 'other side' of morality. He is now in an important sense 'beyond good and evil':

> On that occasion, sitting at tea and chatting with the crew, for the first time in my life I clearly formulated the following for myself: I have neither the feeling nor the knowledge of good and evil, but good and evil really do not exist (and this

[210] Wolman, 1987, p. 44.

[211] Dostoevsky, 1936, p. 704.

[212] *Ibid*., 734.

[213] It is worth noting that Stavrogin was very close to becoming an amoralist who develops a taste for the most unnatural violence (e.g., serial killers). On the next day after rape, and before the suicide, he recalls thinking: "I began to hate her [Matryosha] so that I decided to kill her. […] On my way I kept imagining myself in the act of killing and defiling her." (708). For the role of imagination in development of pathological cruelty see Giannangelo research on the psychopathology of serial killers (1996).

pleased me) and are but a prejudice.[214]

It is crucial for the understanding of the roots of Stavrogin's moral deficiency to find a suitable interpretation of this most intriguing admission. Stavrogin claims that he has no 'knowledge' of good and evil, but it would be a mistake to conclude that he is now unable to make a formal distinction between the two, i.e., that he *forgot* the difference. This, we may safely assume, is highly unlikely given that his general intelligence and memory remained intact.[215] He must have lost moral 'knowledge' in some other, non-cognitive sense of the word 'knowledge'. If that is true, moral competence requires *more* than the ability to (correctly) recognize instances of right and wrong actions. Genuine knowledge of right and wrong, good and evil, and other moral concepts involves (in addition to mastering the purely cognitive content of the concept) having stable emotional dispositions of a certain kind.[216] In a letter to Darya, written days before his suicide, Stavrogin says: "I am still capable, as I always was, of desiring to do something good, and of feeling pleasure from it; at the same time, I desire evil and feel pleasure from that too. But both feelings are always too petty, and never very strong."[217] It is clear from his words that he is at least able to identify good and evil (one cannot desire to do something good unless one has a general understanding of what good is[218]), but yet he lacks that usual emotional component, a *feeling* of good and evil, and this lack places him outside morality. More specifically, the normal emotional dispositions,

---

[214] *Ibid*., p. 712.

[215] For an argument to this effect see "On Forgetting the Difference between Right and Wrong." G. Ryle, 1958.

[216] This is not yet an endorsement of the motivational *internalism* (where moral concepts are necessary or conceptually connected with motivation), but rather a statement of empirical fact that is true of most people, and that even motivational *externalists* can accept.

[217] *Ibid*., p. 685.

[218] The context of his words makes it clear that Stavrogin has *moral* good in mind in this passage.

namely the positive (in the most general sense) conative state in the presence of 'good', and the negative attitude in the presence of 'evil', are not operative in Stavrogin's soul. The association between the morally wrong (evil) act and the feeling of displeasure (that is very strong in most people) has broken down in Stavrogin, or has never taken hold at all. It is in this sense that Stavrogin has no 'knowledge' of good and evil.

Colin McGinn, in his discussion of amoral characters both in real life and in fiction, defines the inherently evil persons in terms of perverse hedonic dispositions:

> Evil will be *either* taking pleasure in pain and pain in pleasure *or* being indifferent to pain. In the case of indifference there is nothing in it for me in the pain of another; his pain simply fails to engage my hedonic dispositions. It may be that such indifference is the end result of evil in the narrower sense, as when a person becomes sated with the pleasure of doing harm.[219]

It is true that an evil character who actively takes pleasure in the pain of another is probably an extremely rare occurrence in real life. "Would any man, who is walking along, tread as willingly on another's gouty toes, whom he has no quarrel with, as on the hard flint and pavement?"[220] asks Hume rhetorically, suggesting that a person who directly desires the pain of another without any external motive would hardly be a human being at all. But as we have seen, Stavrogin fits this description of an evil or amoral person[221] at least to the degree that he is capable, as he admits, of feeling pleasure from causing pain to others or doing something that is considered immoral by all. We may recall here Stavrogin's reason for marrying mentally retarded Marya Lebyadkin – it 'tickled his nerves' (gave pleasure) *because* it was 'monstrous' in the highest degree. McGinn suggests that taking pleasure in

---

[219] McGinn, C., 1997, p. 67.

[220] Hume, D. (1777), 1962, p. 132.

[221] Amoralism or moral indifference is what gives content to the commonsense notion of evil.

doing harm may ultimately lead to *indifference* to both good and evil. The 'disease of indifference' and boredom are also characteristic of Stavrogin. However, such indifference is not the final result of taking pleasure in doing evil, but rather a temporary stage, a state of sensual satiety and utter boredom, that will eventually lead to the search for new pleasures by doing even greater harm to oneself and to others.

To summarize the main accents of this exposition of Dostoyevsky's literary character, we should reiterate that Stavrogin apparently exhibits a number of features constitutive of a classic psychopath, and his amoral, pathological state is confirmed both by his external actions and his internal beliefs and dispositions. As if having Stavrogin or a very similar personality in mind, Martha Stout gives this description of a sociopath: "Sociopaths cannot love; by definition they do not have higher values, and they almost never feel comfortable in their own skins. They are loveless, amoral, and chronically bored, even the few who become rich and powerful."[222] In his confession Stavrogin admits: "I did not love the one I desired so much, and *I could never love anyone*, and there was nothing here but lust,"[223] referring to his courtship of a young lady Liza in Switzerland. We have already noted his constant boredom which eventually drove him to suicide, and the corresponding 'uneasiness' that made him look for ever new stimulations, regardless of their consequences and moral status. The lack of emotional involvement in his relations with others, absence of inner constraints (as a result of absence of 'higher values'), the perverse hedonic dispositions with regard to good and evil

[222] Stout, 2005, p. 188.

[223] Dostoyevsky, p. 717. [my emphasis]

(the lost feeling or knowledge of good and evil), together with charming personality and undamaged intellect[224] complete the picture of a psychopath in his extreme manifestation.

2. The Selective Amoralist: Adolf Eichmann

A selective amoralist was defined in Chapter One as a person who, without *fully* abandoning the moral point of view, limits severely the *scope* of such a disinterested consideration for the welfare of others as inherently required by this point of view. More specifically, a selective amoralist assumes a non-moral attitude toward the people of a certain race, ethnic group, social class or nationality, while still retaining a genuine moral sentiment toward others, those not included in the outcast group. The particular operative reasons and purported explanatory stories why certain individuals are thus singled out for 'special treatment' will certainly differ from one historical context to another, but they need not be seriously discussed or criticized here. The familiar explanations of why these people fall outside the usual 'protective force' of the moral law in terms of their 'hostile intentions,' 'subhuman nature,' or 'heretical beliefs' are for the most part rhetorical devices that do not capture the real reason for the amoral attitudes. What is important for the purposes this section, however, is establishing the fact of such selective application of moral requirements (the fact which, to be sure, hardly needs an elaborate proof), and (more importantly) providing an argument to the effect that such an attitude is (in typical cases) an instance of amoralism (rather than, say, immoralism), which would justify the inclusion of these individuals in the present discussion.

[224] This last fact is especially stressed by the author at the end of the novel: "At the inquest [after Stavrogin's suicide] our doctors absolutely and emphatically rejected all idea of insanity." (688)

It is fair to say that the selective amoralist is by far the most common and (potentially) the most destructive type of amoralist. Their potential destructive efficacy is, of course, a direct consequence of their relatively great number as compared with other, more exotic types of amoral individuals. While perhaps each of us has met at least once a person exhibiting the main traits of a selective amoralist, and can thus testify to the issue from his own experience, it is more appropriate to turn to a case that is publicly accessible and the one which has been described with sufficient details, both in its external and internal (subjective) manifestations. The case that satisfies these criteria is the case of Adolf Eichmann, a notorious Nazi war criminal, who, after his capture in Argentina, was tried in Jerusalem for crimes against humanity in the year 1961. What makes this case even more appealing is that we have a comprehensive account of the life and the trial of Eichmann written by Hanna Arendt, who, as a philosopher, pays special attention to some of the theoretical issues that interest us here most.[225]

In outline, the main facts from Eichmann's biography can be presented as follows. He joined the S.S. organization in the early 30's, when Hitler's party first came to power. Eichmann was originally in charge of organizing the forced Jewish emigration from Germany and Austria, but after the outbreak of the World War II, and the radical change in attitudes among the top Nazi officials toward the "Jewish question," he was placed at the head of the GESTAPO Section IV and was now responsible for the implementation of the new official policy of the "Final solution." This new policy involved transportation of the European Jews to the Eastern 'settlements' in the recently occupied lands, in most cases in Poland. The 'settlements' were extermination camps like the one at Auschwitz, with the capacity to kill tens of thousands of people a day in gas chambers and in mobile gas vans.

[225] Arendt, Hanna. *Eichmann in Jerusalem: a Report on the Banality of Evil*. 1963.

Eichmann was enthusiastic about his new assignment till the very last days of the war; he traveled all over Europe (16 countries in total), and was very efficient in insuring "a steady supply of trainload of Jews" to the killing centers in the East. It is estimated that more than two million people were sent to their death in Auschwitz and other death camps as a direct result of Eichmann's activities.

When atrocities are committed on such a scale the question of the mental health of those directly responsible for the crimes often comes to the forefront. Eichmann was closely examined by a group of psychiatrists during his trial in Jerusalem, and, as Hanna Arendt reports, "half a dozen psychiatrists had certified him as 'normal.' One had found that his whole psychological outlook, his attitude toward his wife and children, mother and father, brothers, sisters, and friends, was 'not only normal but most desirable.'"[226] This authoritative verdict is significant as it rules out with a high degree of plausibility that a purely medical explanation of Eichmann's behavior would be required, and, furthermore, as it confirms our decision to place Eichmann in the category of *selective* amoralists. In Eichmann's case, we do not observe a thoroughgoing inability to assume the moral point of view or to experience the morally relevant emotions (such as shame or guilt), as we have witnessed in cases of some psychopaths, or in Stavrogin's case, but rather a highly selective, and ultimately arbitrary, application of the moral attitudes and moral obligations toward a particular group of people. The validity of the moral rules is not denied; but their scope is severely limited.

But how exactly is Eichmann's amoralism manifested? Since our definition of amoralism contains a subjective element (i.e., a claim about the lack of moral *beliefs* and principles, or, alternatively, a claim about *rejection* of morality), looking at Eichmann's immoral *actions*

[226] Arendt, 1963, p. 22.

only is not enough for placing him in the category of amoralists, as opposed to immoralists or persons with weak will-power. All three categories are prone to wrongdoing in various degrees, but the explanation of immoral behavior will differ in each case. Thus, it is important to establish, using the evidence we have, that Eichmann's point of view (at least at the time when the crimes were committed) is best described as a non-moral or amoral perspective. Luckily, the existing detailed records from the trial give us enough material to work with. Many of the questions that were posed to Eichmann during the trial dealt with his beliefs, values and deepest convictions, and his answers, as I shall argue, are most indicative of his amoral condition at the time.

In the official indictment Eichmann was accused of committing crimes against humanity, while acting on purpose and out of base motives. He pleaded 'not guilty', although he acknowledged that his actions were done on purpose and in full understanding of the consequences. He strongly protested, however, that he acted out of 'base motives.' Hanna Arendt observes:

> As for the base motives, he [Eichmann] was perfectly sure that he was not what he called an *innerer Schweinehund,* a dirty bastard in the depth of his heart; and as for his conscience, he remembered perfectly that he would have had a bad conscience only if he had not done what he had been ordered to do – to ship millions of men, women, and children to their death with great zeal and the most meticulous care.[227]

This may strike the reader as a very disturbing admission indeed. We do not have here a case of a person who is overcome by temptation and, having weak will, commits an immoral act; nor is it a case of someone who is aware of the wrongness of his actions from the moral point of view, but chooses to proceed with the act, judging that the benefits of the immoral

[227] *Ibid*., p. 22.

action are greater than those of the alternative way of conduct. Rather, Eichmann's crimes, it appears, were the result of his sincere willingness to follow his conscience, i.e., willingness to fulfill his *moral duty*. It indicates a personality type and a system of beliefs quite different from that of the ordinary criminal.

That the explanation for Eichmann's motivation for his active participation in atrocities of the Holocaust does not fit into the usual patterns of selfish motives is further confirmed by this remarkable testimony on his part:

> Eichmann suddenly declared with great emphasis, during the police examination, that he had lived his whole life according to Kant's moral precepts, and especially according to a Kantian definition of duty. […] To the surprise of everybody, Eichmann came up with an approximately correct definition of the categorical imperative: 'I meant by my remark about Kant that the principle of my will must always be such that it can become the principle of general laws.'[228]

Eichmann declared himself to be a devout Kantian (sic!) all his life, and yet there were perhaps very few other individuals in history who would be responsible for a comparable amount of human suffering.[229] There is no need to doubt Eichmann's sincerity here in order to explain this apparent paradox. Arendt suggests that Eichmann (as many other Nazi's at the time) was actually following a slightly *modified* version of the Categorical imperative rather than the one formulated by Kant. Without being fully aware of it, "[Eichmann] had distorted it to read: Act as if the principle of your actions were the same as

---

[228] *Ibid*., p. 121.

[229] Yet there is another sense in which Eichmann can be seen as a *true* Kantian – without double quotes. Kant's reiterates in several places that the ban on resisting any supreme lawmaking power is *absolute*, no matter how tyrannical it may become. The following quote from the essay "What is Enlightenment?" is just one of the many examples that could be cited: "It would be ruinous if an officer, receiving an order from his superiors, wanted while on duty to engage openly in subtle reasoning about its appropriateness or utility; he must obey." (1996, p. 19)

that of the legislator or of the law of the land."[230] Since Hitler was the only 'legislator' at the time, and since his decrees were the only 'laws of the land,' it is easy to see how a good 'Kantian', following *this* version of the Categorical imperative, will see it as his *duty* to help exterminating the Jews, Gypsies, homosexuals, Communists, mentally retarded people, and so on:

> All that is left of Kant's spirit is the demand that a man do more than obey the law, that he go beyond the mere call of obedience and identify his own will with the principle behind the law – the source from which the law sprang. In Kant's philosophy, that source was practical reason; in Eichmann's household use of him, it was the will of the Führer. Much of the horribly painstaking thoroughness in the execution of the 'Final Solution' – a thoroughness that usually strikes the observer as typically German – can be traced to the odd notion, indeed very common in Germany, that to be law-abiding means not merely obey the laws but to act as though one were the legislator of the laws that one obeys. Hence the conviction that nothing less than going beyond the call of duty will do.[231]

The Kantian moral tradition (which, we might suppose, Eichmann acquired almost by the process of osmosis since his early childhood) with its strong emphasis on the notion of duty, and on the juxtaposition between one's moral duties and one's inclinations, has only aggravated the overall situation. Now that the core of Kant's morality, namely its requirement to respect humanity in each and every rational being, has been lost and replaced with a formalistic injunction to identify one's will with the *existing* law, one's ability to resist all natural inclinations that might go against the law that sanctions mass murder is seen as the expression of the highest moral virtue:

> There is not a slightest doubt that in one respect Eichmann did indeed follow

[230] *Ibid*.

[231] *Ibid*., p. 121-2.

> Kant's precepts: a law was a law, there could be no exceptions. In Jerusalem he admitted only to two such exceptions: he had helped his half-Jewish cousin, and a Jewish couple in Vienna for whom his uncle has intervened. This inconsistency has still made him feel somewhat uncomfortable […] This uncompromising attitude toward the performance of his murderous duties damned him in the eyes of the judges more than anything else that was comprehensible, but in his own eyes it was precisely what justified him, as it had once silenced whatever conscience he might have had left. No exceptions – this was the proof that he had always acted against his 'inclinations,' whether they were sentimental or inspired by interest, that he had always done his 'duty.'[232]

In what sense, however, is Eichmann's mistaken interpretation of Kant's Categorical imperative a manifestation of *amoralism*? In his detailed typology of immorality Ronald Milo separates a case of *perverse wickedness*, which he defines as a deliberate doing of something that is morally wrong without believing that it is morally wrong. The agent fails to believe that his acts are wrong because they actually confirm to the moral principles that he recognizes as valid; it is just that his moral principles happen to be mistaken or perverse. Such an agent, Milo writes, has at least one extenuating ground:

> Insofar as we conceive of wickedness as perverse wickedness, we have no reason to suppose that the agent of a wicked act is lacking in conscientiousness. Indeed, he may be motivated to act as he does by a sense of duty; for he may believe that one is required to act as he does.[233]

The problem of a perversely wicked person, according to Milo, is that he happens to have bad moral principles. In Eichmann's case, we can say that the action-guiding principle of morality that he was conscientiously following when ordering Jews to be sent to Auschwitz was that modified Categorical Imperative that was the rule of the day: "Act in such a way

---

[232] *Ibid*., p. 122.

[233] Milo, 1984, p. 29.

that the Fuhrer, if he knew your action, would approve it."[234] Why cannot we say then that Eichmann simply had a different *morality*, the one we disapprove of, to be sure, but a morality nonetheless? Why should we classify him and others like him as amoralists, i.e., as persons lacking morality?[235]

The answer would refer us back to the discussion of the nature of morality and moral principles in Chapter Two. Taking up the problem of demarcation between the moral and non-moral spheres, I have argued for the truth of descriptivism, which implies that one cannot have bad (basic or non-derivative) *moral* principles or bad *moral* values[236] since the meaning of 'moral' puts both formal and material constraints on the proper ascription of this term. Thus, if it can be shown that the guiding principle of Eichmann's actions fails to satisfy the conditions discussed earlier, we can legitimately say that he had no guiding *moral* principles (at least with respect to his actions toward his victims), and thus that he was (at the time) in an important sense outside the moral realm – an amoralist rather than an immoralist or a weak-willed person.[237]

---

[234] Attributed to Hitler's personal legal adviser Hans Frank; quoted by Arendt, *op. cit*., p. 121.

[235] Hare considers the case of a fanatical Nazi who is so devoted to the ideal of purely Aryan society that he prefers the realization of this ideal, no matter what the cost in terms of suffering of others. His belief that he ought to exterminate the Jews in order to realize this ideal may be perverse, but it is still, Hare suggests, a *moral* belief. (Hare, 1963 [Freedom and Reason], pp. 165-66.)

[236] To be sure, one *can* have genuine moral principles or moral values that are different from our own, and the ones that we might consider mistaken or misleading. But further analysis will usually reveal a considerable degree of agreement about the *basic* moral principles, and the differences will be largely due to non-moral factual beliefs. When I deny the existence of 'bad' moral principles in this context I refer to a kind of principle that would sanction or permit the kind of actions that are *incompatible* with the meaning of morality (e.g., "Always count the welfare of group X as more important than the welfare of anyone else"). These are not simply *bad* moral principles – these are not moral principles at all.

[237] Aristotle compares a vicious (wicked) person who has perverted action-guiding principles to a city that has bad laws (*Nic. Ethics*, Book VII, Ch. 10). If my argument about the constraints on morality is cogent, at least one interpretation of Aristotle's analogy does not hold: *any* prescription or prohibition can become a law (when

But we *can* show that the principles "Act as if the principle of your actions were the same as that of the legislator or of the law of the land" or "Act in such a way that the Führer, if he knew your action, would approve it" fail to satisfy the necessary conditions for being moral principles. In both versions of this imperative only the *source* of the allegedly authoritative prescriptions is mentioned, but nothing is said about the content of those actions that would or could be approved by the legislator or the *reason* for doing them. The imperative, when taken literally, requires that *any* action-guide be accepted as valid as long as it comes from the proper authority, even if the principles of that action-guide are neither universalizable nor relevant to the well-being of the sentient creatures in any direct way (e.g., the command "Always wear brown shirts"). It does, in effect, advocate a kind of *neutralism* with respect to morality, and to the extent that neutralism was shown to be implausible in Chapter Two (and the argument was made for the truth of *descriptivism*), the action-guide that takes the *will* of the authority (be it society, Führer or God) as the source of moral values, cannot be considered a system of *morality*.

To the extent that Eichmann accepted this modified 'Kantian' imperative, he did accept a non-moral action-guide, and thus can be properly described as amoralist. Yet his amoralism was not due to moral ignorance,[238] but it appears to be a matter of voluntary choice. At some point Eichmann failed to *care enough* about the moral evaluation of his actions, but preferred other non-moral values instead, e.g., the approving attitude of the superiors, possibilities for promotion, etc. Yet unlike the psychopaths or extreme amoralists he was never *fully* deprived

---

properly formulated in universal terms); but it is *not* true that any action-guiding principle can be a *moral* principle (e.g., a principle of ethical egoism cannot).

[238] He was not deprived of the normal upbringing with moral instruction and the emphasis on the importance of moral values.

of the attitude of disinterested concern required by morality. The evidence to his "most desirable" attitude towards his friends and family has already been cited. Arendt points out that Eichmann's conscience functioned normally at times, but there was still one conspicuous peculiarity about him:

> Even during those weeks when Eichmann's conscience functioned normally, it did its work within rather odd limits. […] His conscience rebelled not at the idea of murder but at the idea of German Jews being murdered. It was the same with conscience of a certain Wilhelm Kube, and old Party member and *Generalkomissar* of occupied Russia, who was outraged when German Jews arrived at Minsk for "special treatment" [in gas chambers]. Kube's words may give us an idea of what went on in Eichmann's head during the time he was plagued by his conscience: 'I am certainly tough and I am ready to help solve the Jewish question,' Kube wrote to his superior in December, 1941, 'but people who come from our own cultural milieu are certainly something else than the native animalized hordes.'[239]

The scope of Eichmann's moral concern was strictly limited to the members of a particular group (with the boundaries of this group being somewhat undetermined and depending on the political situation). This attitude of selective application of moral norms was perfectly in accord with the general prescription to the SS members made by Himmler (quoted in full earlier), who was Eichmann's immediate superior at that time:

> It is absolutely wrong to project our own harmless soul with its deep feelings, our kindheartedness, our idealism, upon alien peoples. […] One principle must be absolute for the SS man: we must be honest, decent, loyal and friendly to members of our blood and to no one else.[240]

Kindheartedness, honesty and loyalty are all praiseworthy moral virtues. The claim is made, however, that the behavior sanctioned by these traits of character must be directed

---

[239] Arendt, *op. cit*., p. 90-91.

[240] Quoted in Bennett, J. (1974), pp. 127-8.

toward a selected group of people only. Moreover, it would be *wrong* (*morally* wrong, one must suppose), to exhibit kindness and loyalty to anyone else. A normative system that (implicitly) contains the principle "It is wrong *not* to kill (rob, humiliate, abuse) a member of another ethnic (religious, racial) group" stands as far from being a system of morality as it possibly could. A compliance with such a system presupposes the lack of moral concern or indifference toward the core values of any justifiable moral code.

The example of Adolf Eichmann illustrates a different dimension of amoralism. Unlike the extreme type of amoralist, Eichmann still retains the genuine moral attitude toward a limited group of people – the target-group (e.g., his relatives, members of the same ethnicity, etc.), and yet his arbitrary exclusion of the rest of humanity from the scope of moral concern allows us to view him as belonging to the same general category. Eichmann exhibits many the features indicative of an amoralist in his behavior toward the non-target group even if motivation and justification of his behavior differs significantly from that of the real-life psychopaths or the fictional character of Nikolay Stavrogin. In all these cases the immediate motivating factor (whether it was pleasure, an attempt to escape from boredom, or else an unconditional respect for authority) routinely outweighed the force of the moral reasons – if such reasons were ever accessible for these agents. This radical shift in motivational priorities (or in hierarchy of values) has been described earlier as being abnormal or pathological. The existing overarching similarities in behavior and the intentional structures of a higher level between the extreme and selective amoralist justifies the quest for the *common root* of their respective abnormal conditions. In the following chapters I intend to address this question.

# CHAPTER IV

## THE ROLE OF EMOTIONS IN MORAL PERFORMANCE

> Justice and kindness are no mere abstract terms, no mere moral conceptions framed by the understanding, but true affections of the heart enlightened by reason, the natural outcome of our primitive affections.
>
> Jean-Jacques Rousseau, *Emile*, Book IV.

The distinction between the 'head' and the 'heart' has been part of the philosophical and popular discussions since at least the time of Plato. On the classic account, the two 'organs' symbolically represent the two main capacities of human beings – the capacities for rational thinking and emotional experience respectively.[241] Moreover, it is part of this traditional understanding that there is a frequent tension between these two capacities, and the desires produced by the emotive part may come in conflict with the judgments arrived at by means of pure contemplation. As a result, various attempts were made to eliminate or minimize the potential conflict between the cognitive and the affective parts by overemphasizing the importance of one side, and demeaning the value of the other.[242]

---

[241] Alternatively, *liver* rather than a heart has been identified by some ancient authors as the true seat of emotions.

[242] Emphasizing the importance of reason as opposed to emotions has been perhaps the most common strategy in Western philosophy. But, of course, various 'romantic' trends, which would elevate the status of feelings over that of reasoning, have always been present among the mainstream patterns of thinking as well.

No matter how natural it might feel to us to make a distinction between these two aspects of human psychology, this 'great divide' cannot be treated as self-evident. We should first explore whether there is any real basis for such a sharp distinction between emotions and reason (as popular understanding seems to suggest), and, further, whether the emotional side plays a role in cognitive processes such as attention, memory and logical reasoning – all of which are indispensable for moral performance.

## 1. The Importance of Emotions in Ethics: A Historical Overview

The moral significance of emotions was long ago recognized by Aristotle. Characteristically, Aristotle does not view the emotive human nature as a hindrance for a virtuous life, but rather argues that the task of ethics is to teach a man how to cultivate and control the 'given' natural feelings and inclinations. In his *Ethics* Aristotle conceives of virtues of character as, primarily, dispositions with regard to *feelings*.[243] A virtuous person does not rid himself of emotions but rather trains them through habituation, so that these 'tamed' emotive states would produce virtuous conduct. Indeed, a number of virtuous actions are defined by Aristotle as manifesting the cultivated *emotion.* An act of bravery, for example, is a behavioral response to the well-controlled *feeling* of fear.[244] Thus, promoting the right sort of emotional sensitivity is in many ways crucial, for Aristotle, for a virtuous life.[245]

---

[243] *Nicomachean Ethics*, 1106a 15.

[244] *Ibid.*, 1107a1-30 ff.

[245] Nancy Sherman gives a particularly instructive interpretation of Aristotle's view on the role of emotions in *The Fabric of Character: Aristotle's Theory of Virtue*. Oxford, 1989.

The philosophical interest in 'passions' and their relation to moral life was renewed in the 18th century. When Jean-Jacques Rousseau calls justice and kindness the "true affections of the heart" and "the natural outcome of our primitive affections,"[246] he refers to the innate characteristics which are allegedly part of the human nature, and which, along with the various cognitive potentialities, include capacity and, indeed, strong predisposition for sympathetic concern for other fellow humans in distress – a critical disposition for a moral agent. Rousseau's emphasis on the non-cognitive roots of morality ("primitive affections") was shared by his Scottish contemporaries Francis Hutcheson, Adam Smith and especially David Hume, who placed the notion of a sentiment at the heart of his moral theory.[247] On Hume's account the ability to sympathize with the pain and sufferings of others is indispensable for moral agency, and any person lacking in such ability (e.g., due to some genetic defect) will be forever cut off from this important aspect of social existence.

This anti-rationalist and (eventually) anti-Kantian attitude was further developed by a number of German Romanticists and finally taken up by Arthur Schopenhauer in his important treatise "On the Basis of Morality" (1841). In this work he presents feeling of compassion (*Mitleid*) as the fundamental moral impulse present in most humans to some degree. Compassion, the prerequisite emotion for moral performance, is an innate capacity and cannot be learned: "The mind is enlightened; heart remains unimproved."[248] Schopenhauer admits that there are some rare individuals who lack compassion altogether

---

[246] "Justice and kindness are no mere abstract terms, no mere moral conceptions framed by the understanding, but true affections of the heart enlightened by reason, the natural outcome of our primitive affections." *Emile*, Book IV.

[247] "The hypothesis which we embrace is plain: It maintains that morality is determined by sentiment." (Hume, D. (1777), 1966, p. 289.). Although Hume would disagree with Rousseau about the 'natural' origin of the virtue of *justice* and respect for property.

[248] Schopenhauer, A. (1841), 1995, p. 48.

and exhibit complete egoism and even "pure malice" (i.e., impulse to seek another's harm) in their behavior. In those rare cases, he argues, any reasoning or admonition will be powerless to change their attitudes since the opposite *moral* impulse cannot be created at will.[249]

But it was Max Scheler who gave the discussion of the role of sensuous human nature in moral life a new impetus near the beginning of the 20th century. Through the systematic analysis of the function of emotions in "The Nature of Sympathy" he attempts to securely ground ethics in love and sympathetic concern rather than in reason, with intellectual rigor and ambitiousness that are reminiscent of Kant's attempt to do the opposite.[250] The central objective of Scheler's project was to re-emphasize the significance of emotions and feeling for moral agency and judgment. His main thesis in this work is that there are two fundamental 'spiritual' emotions, love and sympathy, underlying moral action. Sympathy is a *reactive* emotion (i.e., it is the way we respond to the distress of others), whereas spontaneous acts of love have more positive and independent function. It is through acts of love that we experience the *value* of a person *qua* person. For this reason, love is primary in ethical importance as well as in the order of moral performance: "all sympathy is based upon love of some sort and vanishes when love is altogether absent."[251] Love, for Scheler, is a metaphysical principle that is prior to all other emotions; it is what *grounds* morality, and it fulfils largely the same function in Scheler's ethical theory as practical reason did for Kant – love is the discoverer of value and dignity of human beings.

An account similar in its basic lines to some of the classic 'emotivist' conceptions of morality is accepted almost by default by a number of contemporary philosophers as well.

---

[249] Schopenhauer, A. (1841), 1995.
[250] Indeed, A. Vetlesen calls Scheller's project the "inverted Kantianism" (1994, p. 142).

[251] Scheler, M. 1954, p. 147.

Thus, Bernard Williams takes it for granted (even if, as he concedes, it cannot be conclusively demonstrated) that a developed emotional apparatus is a necessary condition for being part of the 'world of morality':

> The man with the extended sympathies, the ability to think about the needs of people beyond his own immediate involvement, is recognizably [within morality]. It does not follow from this that having sympathetic concern for others is a necessary condition of being in the world of morality, that the way sketched is the only way 'into morality.' It does not follow from what has been said; *but it is true*…[252]

Even though many more recent proponents of this view can be cited,[253] it seems sufficient for our purposes to assess the strength of this theory presented in the most general outline without concentrating on any of its specific formulations. What this family of views has in common is the belief that, in Jesse Prinz's words, "any theory of moral development and moral behavior that emphasizes reasoning alone leaves something out."[254] The non-cognitive, emotive component, on this view, is an essential part of a description of moral life and moral understanding. Moral judgments (as well as choices based on those judgments) are inextricably bound to moral emotions, and cannot be grasped in isolation from them. The systematic failure in moral judgment and moral performance should in turn be explained (at

---

[252] Williams, 1972, p. 14, (my emphasis).

[253] E.g., Duff, 1977; Blum, 1980; Slote, 2001; Prinz, 2005; Eisenberg, 2005. Although all of these authors recognize the necessity of the emotive component for a developed moral agency they may differ in their accounts of the relevant importance of other factors as well. The positions may range from the emphasis on the exclusive and central role of empathy for moral behavior to a more modest view of Eisenberg, who argues that "emotional factors are *at least as important* as purely rational ones in explaining moral motivation and behavior." (p. 75). Needless to say, emphasis on the affective aspect of moral experience is very important for various feminist approaches to moral knowledge and performance (e.g., M. Little, 1995).

[254] Prinz, J., 2005, p. 1.

least partly) by the agent's reduced capacity for empathy, shame, feeling of guilt or any other set of the morally relevant emotions.[255]

Before exploring the explanatory alternative of amoralism along the lines sketched above it is prerequisite, however, to address the question of the *nature* of emotions first. But since it is impossible in the present context to develop a full-blown theory of emotions and do full justice to all objections and alternative views, the discussion below will be of necessity highly selective in its justification and somewhat programmatic in its positive account of the emotive phenomena.

## 2. What Emotions Are

To begin with, it is obvious that we *do* experience emotions, feelings and moods which phenomenologically appear to be quite different from the various cognitive processes, such as contemplating, memorizing and reasoning. One notable difference is that many basic emotive states (e.g., fear, surprise, anger) are often forced upon us by the outside world, whereas we are usually in control of our cognitive processes (e.g., one can actively *recall* a poem, concentrate *attention* on the matter at hand, etc.). On the other hand, these two mental capacities are not altogether independent. Especially with the more complex emotive states (e.g., guilt, pride, envy) some conscious cognitive processing becomes necessary for experiencing the emotion in question.[256] One would not feel guilty, for instance, unless there is an understanding of what is expected from the point of view of some normative system, and the recognition that one falls short from this expectation. Likewise, it is common

[255] For example, Antony Duff argues that "Psychopathy involves an incapacity for such emotional and moral responses as love, remorse and concern for others." (1977, p. 191).
[256] See Ekman, Paul. 1992.

knowledge that the various cognitive functions, such as practical reasoning that leads to intentional actions are frequently affected by the emotive states of the agent – in many cases, without our conscious awareness of such an influence.[257]

Given this apparent interdependency between cognitive and affective sides of the human psyche, it would be misleading in many cases to ask whether a certain form of behavior has 'emotive' or 'cognitive' origin, where this disjunction is taken in an exclusive sense. To be sure, we could produce paradigm examples of the purely non-cognitive reactions, such as reflex-like feelings of fear or disgust, and, similarly, a number of simple operations with clearly defined rules, such as counting, might exemplify the workings of the cognitive faculty with no or little 'emotional mixture.' But this strict dichotomy seems artificial in a variety of other, more complex (and more interesting) contexts. The aesthetic response to a work of art such as a picture (a *prima facie* non-cognitive reaction) requires (among other things) *noticing* the wavy lines in the middle of the canvas, *recognizing* a person or a place depicted, *remembering* various details of the picture, *interpreting* the symbolism of the artist, *comparing* it with other works of art one has seen, and a number of other cognitive processes. Indeed, the *admiration* that a picture inspires seems to be the supervening result of all the cognitive processes that were involved in a picture's evaluation. To admire the work of art *is* to notice, recognize, remember, compare, interpret the many features of this work. This dependency is evidenced by the fact that a desired aesthetic response can sometimes be induced in an otherwise insensitive viewer by explaining,

[257] A fact that gives rise to a popular speculation that an *emotionless* creature (e.g., Mr. Spock from the *Star Treck* series) would exhibit superior intellectual qualities as compared with the emotional, passionate creatures like ourselves. This picture is challenged by Dylan Evans, who argues in his recent book for the thesis that "a creature who lacked emotions would not just be less *intelligent* than we are; it would be less *rational* too." (Evans, D., 2001, p. 180).

pointing to, and emphasizing the relevant details of the artwork. The exclamation "*now* I see it!" often rewards such an effort.

But before we can fully appreciate the significance of this insight for moral (and amoral) performance, there is one theoretical possibility that needs to be dealt with before we can proceed any further. Namely, we should consider the view that denies the existence of emotions as a separate class of mental phenomena altogether.

## 2.1 What Emotions Are Not: The Weakness of the Cognitivist Approach

The apparent dependency of some emotive reactions on cognitive processes[258] and the resulting problem of demarcation between the two kinds of mental events led some thinkers toward the reductionist account which tries to explain *all* emotive phenomena in terms of entertaining certain thoughts, having beliefs or making judgments. To experience an emotion, on this view, is identical to making a judgment about the world. As a result, this view denies that there is anything unique to experiencing an emotion as opposed to undergoing some cognitive process. In effect, it eliminates all emotive phenomena from the list of the philosophically interesting objects of study, and invites us to shift attention to the cognitive side instead. We may refer to this line of thought as the Propositional Attitude theory of emotions. Since one possible explanatory account of amoralism blames the emotive *deficiency* on the part of amoral individuals, we should first inquire whether there is a distinct class of mental phenomena denoted by the term 'emotion' that one can lack or be deficient of in the first place.

[258] Indeed, the dependency is mutual. One's emotive state will influence one's ability to reason, pay attention, memorize the poem, etc.

In its most straightforward form the Propositional Attitude theory is defended by Robert Solomon when he makes the following identity claim: "My anger *is* that set of judgments […]; an emotion is an evaluative (or normative) judgment."[259] And again, using other instances of emotion as examples, he writes:

> My embarrassment is my judgment to the effect that I am in an exceedingly awkward situation […]; my sadness, my sorrow and my grief are judgments of various severity to the effect that I have suffered a loss.[260]

The view that eliminates emotions altogether as a separate class of mental states immediately strikes one as rather implausible and outright paradoxical. It clearly goes against what most people believe about this subject, which, in turn, incurs a burden of accounting for the 'mistaken' folk-psychological conception of emotions on anyone defending this view. Besides, it is not immediately clear what deeper purpose may be achieved by such a taxing move. Yet, on a second look, it appears that by denying emotions a special place in the taxonomy of mental occurrences, one may harmonize a particular conception of human nature with the demands of moral responsibility. The need to incorporate the notion of 'radical responsibility' with the phenomenological fact of affective human nature is what really seems to motivate the Propositional account of emotions.

One platitude that is part of the folk-psychological account is that emotions are essentially passive occurrences that a person undergoes or suffers, rather than intentionally produces. Furthermore, many of the morally relevant actions seem to be the direct outcome of the emotive state of a person on a given moment. A person experiencing extreme rage will most likely act in a way that affects the welfare of others. Yet, to the extent that his actions

---

[259] Solomon, 1977, p. 185. (emphasis in the original). The original formulation of this theory appears much earlier in Anthony Kenny's *Action, Emotion, and Will* (1963), but a very similar version of this view is defended already by Philippa Foot in her "Moral Beliefs" (1959).

[260] *Ibid*., p. 186.

were directly caused by this emotion (rage), over which the agent had no direct control, it seems counterintuitive to hold him fully responsible for his behavior. Similarly, a person failing to act properly in a morally relevant situation may be partly excused if he is 'by nature' less sensitive than an average person, and is not easily prone to pity or other similar feeling.

The propositional account, on the other hand, by re-interpreting emotions as judgments, eliminates references to the allegedly passive nature of emotive experience, and thus, does away with a whole class of justifications which seek to mitigate the blameworthiness of one's actions. Robert Solomon's thesis that having an emotion is something we choose to 'do' rather than merely suffer brings new acuity to the notion of moral responsibility. If Solomon (following Sartre[261]) is right, and "we make ourselves angry, make ourselves depressed, make ourselves fall in love,"[262] then no reason is left for judging the immoral actions of an enraged, jealous or depressed person with lesser severity than the actions of a more composed wrongdoer. The domain of human freedom is thus extended to the area traditionally seen as the seat of irrational impulses and uncontrollable reactions – it is indeed *radicalized* in Sartre's original sense.

Nonetheless, despite its attractive consequences for the existentialist's agenda, the Propositional theory of emotions must be rejected. In what follows I shall list several considerations that make any reductionist cognitivist account of emotions highly implausible.

---

[261] See Sartre, J.- P., *Sketch for a Theory of the Emotions*. London: Methuen, 1971.

[262] Solomon, 1977, p. 185.

First, as Paul Griffith points out,[263] there are clinical instances of 'objectless' emotions such as depression, elation and anxiety. In those cases, the emotions in question involve no intentional object (they are not *about* anything specific in the world), and thus involve no propositional attitude. Strong emotional responses to *imagined* objects are also hard to explain along the cognitivist lines.[264] Secondly, various *reflex* emotions, such as fear and disgust cannot be correlated with making a judgment about the world in the conventional sense. According to the cognitivist view of emotions defended by Sartre and Solomon, the fear of, say, spiders must be interpreted as an evaluative judgment that spiders are dangerous. However, there are well-documented cases where the reflexive fear of spiders co-occurs with the (sincere) conviction that spiders are harmless. Certain immediate emotive reactions seem to be (in Griffith's phrase) 'informationally encapsulated,' and take place despite and independently of all the relevant judgments.[265] Thirdly, since the cognitivists identify the evaluative judgment (e.g., 'x is dangerous') with the content of a specific type of emotion (e.g., fear), they can be accused of inflating the number of token-occurrences of emotions beyond necessity. Many drug-users believe drugs are dangerous but continue using drugs without fear.

Finally, and most importantly, the propositional theory of emotions fails to account for the impressive amount of evidence that correlates all of the so-called 'basic' emotions

---

[263] Griffith, 1997, p. 28.

[264] Greenspan (1988), among others, argues that people can have full-blown emotional responses to imagined situations. These reactions to imagined objects are in no sense less 'genuine' than the ones evoked by the real world.

[265] Socrates might have had something like the cognitivist account of emotions in mind when he argued that one will cease to *fear* death once he realizes that death is one of two things, the permanent state of unconsciousness or the conscious afterlife, both of which are pleasant in their own right (*Apology*, 29b; 40d). Common experience suggests, however, that fear of death is largely independent of one's beliefs about the postmortem state.

(e.g., surprise, fear, sadness, anger *et al.*)[266] with specific physiological changes, such as facial expressions, expressive vocal changes, and changes in endocrine and nervous system (e.g., adrenaline release and increased heart rate).[267] The cognitivist theory leaves this stable connection (which does not depend on a cultural context) between a cognitive process (making of an evaluative judgment) and physiological changes totally unexplained. It remains a mystery, on this view, why some judgments trigger such dramatic physiological reactions, while other judgments with similar content leave the person unaffected.

In view of the objections mentioned above, the radical reductivist approach to emotions cannot be sustained. It remains then to give a brief outline of the alternative approach to affective phenomena, which does not seek to eliminate emotions altogether.

2.2 A Programmatic Account of Emotions

In light of the serious problems of the eliminativist strategy of the propositional theory, I shall maintain that emotions are in an important sense real (i.e., they cannot be reduced to the purely cognitive processes, and the term refers to a distinct psychological class of events), and as such, play an important independent role in determining human behavior.

Our vernacular use of the concept of 'emotion' somewhat indiscriminately covers a range of psychological phenomena that, on a closer look, exhibit significant idiosyncrasy within this folk-psychological category, and thus warrant a more sensitive classification. Many emotion researchers believe that some emotions are more fundamental or basic than others, and should be treated differently.[268] The standard list of 'basic emotions' includes

[266] More on 'basic emotions' in the next section.

[267] Ekman, 1992; Ekman and Friesen, 1971.

[268] Ekman, 1992; Griffith, 1997; Evans, 2001; Prinz, 2005. The six basic emotions admit of degrees, i.e., may vary in intensity, which is reflected in a rich emotion-concept vocabulary that exist in all languages. Thus sadness may range from slight pensiveness to grief; fear – from apprehension to extreme terror; surprise – from brief distraction to amazement, etc.

anger, fear, disgust, surprise, happiness (joy) and sadness.[269] All six emotive reactions have well-documented cross-cultural facial expressions associated with each emotion, musculoskeletal responses such as flinching, and a series of other coordinated physiological changes. They have undisputed neural basis in the limbic system of the brain, the phylogenetically ancient portion of the cortex which surrounds the brain stem,[270] and their pancultural existence is usually explained (with greater or lesser success) by the various evolutionary advantages they give to an animal who can experience an emotion in question. The descendents of *fearless* proto-humans, or those with indiscriminate eating habits (due to lack of feeling of *disgust*) are not among us today for obvious reasons.

An important thing to notice about these six basic emotions is that they require the bare minimum of cognitive involvement on the part of the agent. Their instances are largely immediate reflexive reactions to the perceived stimuli of the environment. Whatever cognitive processing is needed to appropriate the data of the senses, most of it occurs unconsciously. The basic emotive responses are largely independent of the conscious cognitive processes – the bypass the neural circuits of the outer (phylogenetically more recent) cortex responsible for the higher cognitive functions. As Zajonc argues, this fact of 'autonomy' of basic emotions suggests a certain chronological order in the evolution of the various capacities of the human brain:

> He limbic system that controls emotional reactions was there before we evolved language and our present form of thinking […] When nature has a direct and autonomous mechanism that functions efficiently, it does not make it indirect and dependent on a newly evolved function [cognition]. It is rather more likely

---

[269] With the possible addition of *contempt* as a distinct basic emotive reaction (Ekman and Friesen, 1986).

[270] Zajonc, 1980; Damasio, 1994.

> that the affective system retained its autonomy, relinquishing its exclusive control over behavior slowly and grudgingly.[271]

The six or seven basic emotions can be compared and contrasted with the more complex, cognitively mediated emotive responses, such as jealousy, guilt, love, sympathy, envy and shame. Paul Griffith refers to these states as the 'higher cognitive emotions', and I shall follow this convention here.[272] These non-basic emotions are much more culturally dependent than the basic affects, both in terms of the manner of manifestation and the antecedent causes, and, admittedly, less well understood. The various 'moral' emotions that traditionally have been assigned a central role in moral performance will all belong to this category. The more important it is for us to inquire into the nature of these mental states.

How do higher cognitive emotions relate to the basic ones? One tempting theoretical move is to suggest that all higher emotions are *blends* of several basic emotions, and thus the main task of the researcher is to identify which basic emotive reactions are involved in a given higher cognitive affective state, and in what proportion. Jesse Prinz in his recent article argues for this possibility.[273] He suggests that various non-basic emotions can be generated by blending basic emotions through the process of *calibration*. For example, pride is joy (a basic emotion) calibrated to one's accomplishments.[274]

This reasoning has immediate relevance to the problem of moral performance. According to Prinz, the amoral behavior of a psychopath may be explained by the inability to

---

[271] Zajonc, 1980, p. 169.

[272] Griffith, 1997. An alternative labeling of the two kinds of emotive phenomena is used by Vetlesen (1997), who uses the term 'feeling' to refer to the 'rawer' basic emotions, and reserves the notion of emotion proper for the 'higher' emotional states. Damasio (1994) distinguishes between 'primary' and 'secondary' emotions which seem to be coextensive with these two categories.

[273] Prinz, 2005.

[274] *Ibid*., p. 7.

experience one or two basic emotions which, in turn, has drastic consequences for his capacity to generate various morally relevant non-basic feelings - guilt and shame being the most crucial ones. Prinz interprets guilt as sadness calibrated to situations in which one has caused harm to someone that one cares about, and further analyzes shame as a blend of sadness and aversive self-consciousness.[275] The explanation of amoral behavior follows straightforwardly once we postulate a (genetic) deficiency in sadness – a basic emotion that figures into both moral feelings - in most extreme amoralists. Both guilt and shame are inhibitory emotions that tend to control impulsive behavior in normal individuals. A (postulated) general deficit in sadness in psychopaths prevents them from generating these crucial 'higher' emotions, which in turn results in moral retardation that we captured by the term 'amoralism.'[276]

Still, despite its apparent simplicity and the promising explanatory potential, the blending theory of higher emotions cannot be the full story. Paul Griffith describes several inadequacies of the blending theory, three of which I find to be especially damaging to the view in question.[277] First, many of the higher cognitive emotions (e.g., love, jealousy) are sustained responses, usually lasting for long periods of time, whereas all the basic emotions (the alleged constituents of love, jealousy, etc.) are brief responses. It is not clear how the admixture of several rapid, reflex-like emotive reactions can produce something as stable as many higher emotions are. Secondly, unlike the basic affects, the higher emotions lack stereotypical behavioral and physiological consequences. There appears to be no single facial

[275] 'Aversive self-consciousness' is his original contribution to the usual list of basic emotions. He explains it as "an emotion that arises when one receives an unwanted attention from others" (p. 7).

[276] *Ibid*., pp. 8ff.

[277] Griffith, 1997, pp. 101-103.

expression associated with, say, envy, yet one would expect that an unvarying blend of two or more basic emotions would produce a universally recognizable facial and behavioral response. Finally, the blending theory of emotions fails to account for the evident dependency of the non-basic emotions on cognition. The degree of cognitive involvement in particular cases of emotive reaction may differ, by it seems to be ubiquitously present in all instances of non-basic emotions. Yet, as Griffith observes, "blending several reflex-like responses does not produce something more cognitively involved."[278]

A more plausible account of higher emotions is proposed by Antonio Damasio, who explains them as involving *both* the basic affects and the distinctive higher cognitive activity.[279] This addition of cognitive activity complicates the economical picture of the blending theory, but, nonetheless, seems to better correspond to our deep intuitions about emotions,[280] and avoids many of the problems cited above. Yet the price for this advantage is the precision reduction in the analysis of specific instances of emotions.

We can easily appreciate how difficult it is in many cases to demarcate the sphere of the cognitive from the non-cognitive processes and responses. The difficulty here is not primarily 'technical', but is rather due to the complex nature of the most 'higher' mental functions. Contrasting the cognitive contribution to choice and behavior with the affective influence is a somewhat unnatural and unnecessary task once we get beyond the primitive instinctive reactions and the most basic intellectual operations. The same reasoning would apply within the moral context to actions whose underlying motivating cause is usually

---

[278] *Ibid*., p. 102.

[279] Damasio, 1994; Blum (1980) is proposing a very similar account.

[280] E.g., by retaining basic emotions as *elements* in higher emotions Damasio accounts for a widespread belief that *all* emotions involve *involuntary* arousal at least to some degree.

referred to as *sympathy*. Sympathy is a complex response which clearly involves both cognitive and non-cognitive workings of the human mind, but it is perhaps futile trying to analyze this common reaction into its constitutive parts, assigning a precise relevant weight to the cognitive and affective constituents. Nor does it have to remain the same in all cases. A child of five feels sorry for his sister after witnessing her falling down the stairs and bruising her leg. As a result, he tries to comfort her. A successful businessman sympathizes with the victims of the genocide in Darfur after reading about it in the newspaper, and as a result, donates to the charity working in Africa. There is no doubt that the child's reaction is much more immediate and, we may say, almost purely emotional,[281] whereas the pathway leading to sympathy that a person experiences in the second example is more indirect, requiring greater involvement from the intellectual side. Yet both behavioral responses are motivated by essentially the same higher cognitive emotion of sympathy. Lawrence Blum generalizes from these kinds of common-sense observations to all altruistic emotions when he says:

> Sympathy, compassion, or concern are directed towards others in virtue of their suffering, misery, pain travail. And so the altruistic emotions have a cognitive dimension: the subject of the emotion must regard the object as being in a certain state (e.g., of suffering).[282]

One consequence of this view is that we cannot, strictly speaking, refer to a person lacking in sympathy or compassion as suffering from a *non-cognitive* defect.[283] If 'cognitive dimension' is a necessary constituent of a given emotive state, a purely *cognitive* failure may result in the

---

[281] Flack and de Waal (2000) argue that even non-human primates – especially the great apes – are capable of varying degrees of empathy. If the empathic-like response that apes are capable of is indeed phylogenetically continuous with human empathy, then the level of cognitive sophistication required for experiencing the emotion in question cannot be too high.

[282] Blum, 1980, p. 12.

[283] This would only be possible on the blending theory of higher emotions, which interprets compassion or sympathy as a mixture of 'pure' non-cognitive affective states.

inability to experience sympathy. The sense in which an amoralist might be said to be cognitively deficient will be explored in the following chapter.

To summarize above discussion, the denotation of the term emotion, the way it is used by the competent members of a language community, subsumes at least two different kinds of psychological state: the basic and non-basic emotions. The distinction is not arbitrary, but corresponds to real distinctions in human psychology. The qualification 'non-cognitive' would most fittingly apply to the basic emotive states, which are universal,[284] irruptive, immediate reactions, such as joy or fear, involving little or no references to consciously accessible beliefs and desires. The experience of the cognitively mediated non-basic emotions, on the other hand, partly depends on the cognitive activity on the part of the person, which in turn is sensitive to one's cultural context. The way one falls in love, manifests jealousy or experiences shame will differ from one cultural milieu to another, depending on the beliefs one holds on the variety of subjects. The non-basic emotive states, thus, occupy a strategic middle point between the pure emotive reactions and pure cognitive processes, requiring the coordinated workings of both human functions.

## 3. The Emotive Deficiency Theory

I shall refer to a position that singles out emotive failure as the primary (if not the only) cause of the moral failure, both in judgment and in action, as an 'emotive deficiency theory.'

But a clarification is in order. It is unlikely that an amoralist is incapable of experiencing *all* emotions, both basic and non-basic. Such a thoroughgoing emotive deficit hardly exists among human beings. Presumably, then, his deficiency is less extensive, and he

[284] I.e., pancultural.

lacks capacity (or has a diminished capacity) for experiencing only those emotions that are crucial for moral performance. But *which* emotions are crucial for moral performance? The answer to this question already presupposes that at least some emotions *are* necessary for moral agency. This assumption, of course, is by no means self-evident, and needs further justification. Yet, granting the assumption for the purposes of the argument, what is the list of emotions (we may call them 'moral emotions') that play this determining role in moral performance?

As one can expect, the list will differ from one philosopher to another. We have seen Max Scheler taking *love* (in his special sense of this word) as the most fundamental moral emotion which is more primordial than sympathy or compassion; and Jesse Prinz arguing that *guilt* and *shame* are the two emotive states that function as preventive, inhibitory mechanisms for all moral agents. But it is perhaps fair to say that the most plausible versions of the emotive deficiency theory would identify *empathy* (or one of its cognates and derivatives) as an emotive state that underlies moral behavior. By extension, an amoralist is someone who is unable to experience empathy, or, if empathy admits of degrees, is unable to experience it with intensity that would be sufficient for moral performance. This is the view defended, among others, by Arne Vetlesen when he says:

> According to my explanatory model, an overt failure to recognize and respond in a situation involving another's suffering suggests that the nonaffectedness displayed may stem from the person's diminished or even severely damaged capacity for empathy. […] A lack of empathy proves a sufficient condition of moral blindness.[285]

[285] Vetlesen, 1994, pp. 163; 277. The same idea becomes central in certain feminist approaches to moral performance. Thus Margaret Little writes: "Possession of certain desires and emotions turns out to be a *necessary* condition of discerning moral properties, and hence must form part of even the ideal's observer's epistemic repertoire." (1995, p. 120).

As I take empathy being the most likely candidate for the role of the prototypical moral emotion, in what follows, I shall limit my discussion to this version of the emotive deficiency theory only.

## 3.1 Defining Empathy

'Sympathy', 'sympathetic concern', 'the sentiment of humanity', 'compassion', 'commiseration' and 'empathy' are the most common terms used in connection with the view in question, and, as their similar etymology suggests, they can perhaps be employed interchangeably in most contexts. Since, however, most of the contemporary writers in the field of moral philosophy and psychology give marked preference to the last term (namely, empathy), I shall also follow this convention, and assume that the functional description of its workings in all essential features will *ceteris paribus* apply to the others as well.[286]

The English word 'empathy' dates only to 1903, when Titchener coined it as a translation of the German '*Einfűhlung'*, which literally means "in-feeling."[287] The term is now used in psychology, psychiatry and philosophy, and despite the superficial simplicity and intuitive clearness of the theory of emotional deficiency, it is not at all obvious what exactly empathy involves and how it contributes to the rise of a well-developed moral character. A psychologist Uta Frith gives the following description of the normal operation and preconditions of empathy:

> What empathy requires is the ability to know what the other person thinks or feels

[286] Even those authors who make a technical distinction between empathy and sympathy (e.g., Eisenberg, 2005) take empathy to be more fundamental than any other affective motivation.

[287] Preston and de Waal, 2002.

> despite the fact that it is different from one's own mental state at the time. In empathy one shares emotional reactions to the other person's different state of mind. Empathy presupposes, amongst other things, a recognition of different mental states. It also presupposes that one goes beyond the recognition of difference to adopt the other person's frame of mind with all the consequences of emotional reactions.[288]

In agreement with the above analysis of non-basic emotions, Frith emphasizes that the operation of empathy (even if it is partly a *non-cognitive* reaction) requires the exercise of certain other cognitive abilities - knowledge of mental states of another person and recognition of his emotive condition. Yet, if our criticism of the Propositional Attitude theory is sound, empathy itself cannot be *identified* with any of these cognitive processes or judgments. Frith's description, however, says little more on the basic affective component of this important moral emotion.

A number of similar functional descriptions of empathy as, for instance, "the affective capacity involved in recognizing and being moved by another's distress"[289] emphasize the behavioral outcomes or the identifiable empathy-related *responses*, but are usually too vague and too loosely formulated to give adequate understanding of the limits of application of this concept. If empathy is "involved in recognizing another's distress," does that also imply that one *cannot* recognize a person in pain *unless* one has empathy? In what sense is it 'involved'? A sadist is certainly capable of recognizing a person in distress and be *moved* by such recognition (he does have specific 'emotional reactions', i.e., he may experience satisfaction), but it would be odd to ascribe empathy (in its colloquial understanding) to a sadist who is sensitive to *and* takes pleasure in another's suffering. If, on the other hand, empathy *can* be ascribed to a sadist, then we are forced to look for some other (perhaps,

---

[288] Frith, U. 1989, pp. 144-5.

[289] Kennett, J. 2002, p. 341.

cognitive) cause of the sadist's amoralism, which in effect amounts to abandoning the 'emotive deficiency' theory.

This kind of questions and counterexamples call for a more precise analysis of empathy, if empathy is indeed a distinct emotive state which cannot be fully reduced to any other cognitive or basic non-cognitive human capacity. In a recent empirical study of the role of empathy in moral motivation and behavior Nancy Eisenberg recognizes the need for a more careful definition of this important notion and proposes to define it in the following way:

> We define *empathy* as an affective response that stems from apprehension or comprehension of another's emotional state or condition and is similar to what the other person is feeling or would be expected to feel in the given situation.[290]

Eisenberg's definition has a clear advantage over the earlier attempts in that it rules out the possibility of a sadist having genuine empathy – the emotional responses that a sadist has at a sight of a person's suffering (i.e., pleasure, satisfaction) are *not* similar to the ones that the person himself normally experiences (i.e., anguish, terror). Emotional contagion that takes places in empathy 'transfers' the content of the original emotion. Furthermore, the similarity of emotive reaction ensures not only the motivational efficacy of that state,[291] but also the similar *direction* of this motivating impulse (by contrast, a sadist who does react emotively to the victim's distress is thereby motivated to increase or prolong the suffering – an impulse very different from the victim's). The definition still preserves the essential dependency of the *affective* response on the purely cognitive operations, namely, apprehension or *comprehension* of another's condition. A person who is unable to recognize

[290] Eisenberg, N. 2005, p. 75.

[291] We may agree with Hume that all 'passions' are motivationally efficacious without necessarily accepting the stronger position that only 'passions' have this capacity.

another person's distress due to some defect in his understanding will not be able to experience empathy even if there be no deficiency in his emotive makeup *per se*. A young child easily sympathizes with an adult who exhibits a sad face and emits some characteristic 'pain-sounds', but will remain unmoved if the adult's distress is expressed by words and gestures that are beyond child's comprehension. In that case, the child's 'coldness' will be properly explained by reference to his cognitive limitations rather than to his emotive deficiency.

I have so far interpreted Eisenberg's definition of empathy as an affective response which requires *content similarity* between the two *basic* emotive states (i.e., that of the empathizer and the person whose condition evokes this response), and whose proper operation depends on certain cognitive capabilities. In effect, to be able to empathize is to be able to experience a similar basic emotion (e.g., sadness or degrees thereof, joy, fear, etc.) where this feeling is evoked not by its usual environmental stimuli (e.g., the loss of a loved one, the presence of danger, etc.), but by a recognition of another person's condition. A person who feels sadness because he lost something valuable experiences a basic emotion; a person who (involuntarily) acquires the same feeling[292] as a result of seeing and recognizing sadness on his neighbor's face shows a capacity for a higher cognitive emotion of empathy.

I shall now consider one objection to this construal of empathy as primarily an emotive capacity. It is perhaps this peculiar dependency of empathy (or sympathy) on cognitive recognition of another's emotional state that prompted Richard Boyd, among others, to classify sympathy rather as a primarily *cognitive* mechanism, and argue that an amoralist who lacks sympathy, i.e., "the morally unconcerned person, the person for whom

[292] The same in content, but not necessarily the same in intensity.

moral facts are motivationally irrelevant" probably suffers from a "*cognitive* deficit with respect to moral reasoning."[293] Boyd argues for this conclusion by assuming a crucial premise, namely, that

> [t]he capacity to *recognize* the extent to which others are well or poorly off with respect to the homeostatic cluster of moral goods and the capacity to *anticipate correctly* the probable effect on others' well-being of various counterfactual circumstances – depends upon their capacity for sympathy, their capacity to imagine themselves in the situation of others.[294]

Both *recognition* and *correct anticipation* are the paradigm cases of cognitive abilities. Boyd seems to be arguing here that one is *unable* to recognize another in distress unless he also at the same time *sympathizes* with that person. Thus sympathy becomes a necessary condition for the proper cognitive functioning, and an amoralist deficient in sympathy becomes *ex hypothesi* deficient in cognitive abilities as well. This, however, is a much stronger claim than I am willing to accept.

Two responses are possible here. First, it appears (from the passage quoted) that Boyd defines sympathy as "the capacity to imagine oneself in the situation of others." If *that* is the stipulated definition of sympathy, then it clearly becomes a cognitive capacity. However, I have argued above (with Eisenberg) that sympathetic concern or empathy essentially involve similarity in *emotional* states, and no simple recognition of another's condition can grant this (as the example of a sadist might show).

Secondly, we may deny the premise that having sympathy is *necessary* for being able to perform some other cognitive functions (e.g., recognition of a person's condition). Quite

---

[293] Boyd, R. 1988, p. 213. The same intuition is shared by Arne Vetlesen (1994), who argues that "as far as the moral judgment is concerned, the exercise of judgment *presupposes* and *is made possible* by our 'having' certain emotions." (p. 157).

[294] *Ibid*. [Boyd's emphasis]

the opposite, as I have argued above, the dependency is reverse. One must first be able to recognize the person in pain, and sympathy (a comparable emotive response) may or may not follow. Surprisingly, Boyd himself seems to weaken his previous reasoning by acknowledging that "there could be cognizing systems which are quite capable of assessing moral facts without recourse to anything like sympathy."[295] He later lists sociopaths and con-artists as the most plausible candidates that would satisfy this description in actual life. Once Richard Boyd agrees that it is at least possible to know all the morally relevant facts, and reason correctly about them (as judged against the standards of formal and informal logic), he leaves no further reason to take sympathy as a purely cognitive mechanism, and there is no need to suppose that the lack of sympathy should have those devastating consequences for one's ability to reason correctly about moral facts.

3.2 Autism, Moral Agency and Extreme Amoralism

We may call a motive (a desire) to relieve the sufferings of others or to help others in distress (including the desire to abstain from causing direct harm) the 'moral motive'.[296] We have seen that a cognitive recognition of the presence of pain and suffering in another is (at least) a *necessary condition* for being morally motivated. What is crucial for defenders of the emotive deficiency theory, however, is whether recognizing another's distress (which requires at least rudimentary mastery of the moral and morally relevant concepts) is by itself *sufficient* (in normal individuals) for acquiring a moral motive (of whatever strength), or whether an operation of the intervenient emotional mechanism (say, empathy) is also

[295] *Ibid*., p. 214.

[296] The 'moral motive' thus defined may or may not be part of the explanation why people are sometimes motivated to act *justly*, since a concern for justice (as it was already noticed by Hume) is a somewhat peculiar moral concern, which might have a different origin from the more common moral concerns that could be reduced to benevolence and non-maleficence.

*necessary* for that purpose. If the first alternative is true,[297] we may safely dismiss the emotive deficiency theory and concentrate instead on the search for some cognitive defect in a person who consistently fails to be motivated in the morally appropriate way. If, on the other hand, the second alternative is substantiated, and cognitive recognition of the relevant properties of the world (e.g., pain) is *not* by itself enough for being morally motivated, then it follows that any given case of amoralism (where the lack of moral motivation is seen as a main *symptom* of this condition), may be *either* due to the intellectual limitations of an agent (as in cases of a young child or a mentally retarded person), *or* due to the lack of empathy (as in cases of at least some autistic individuals or Richard Boyd's dispassionate "nonhuman cognizing systems"[298]), or *both*. In other words, the truth of the second alternative would *not* yet warrant the exclusive importance of the emotional factors for a moral life. Indeed, it would leave the question of the primary cause of amoralism in a *specific* case entirely open.

There exists, however, strong evidence, coming from the field of clinical psychology, that may further undermine the claims to exclusive importance of the purely emotive factors for a competent moral agency. The evidence comes from the studies in autism. There is a general consensus among the psychotherapists that autistic individuals are similar to psychopaths in at least this respect: both are severely impaired in their ability to sympathize with the distress of others, to develop connectedness to others and enter into meaningful relationships with other people.[299] A standard description of autism will usually include a list

---

[297] The truth of this alternative would seem to imply the *motivational powers* of practical reason. For a defense and explanation of this Kantian view see David Velleman (1996), *The Possibility of Practical Reason*.

[298] Boyd, R., 1988, p. 214.

[299] It is perhaps significant that the first researchers of this condition in the 1940's have referred to this disorder as 'autistic psychopathy'. (See, for instance, the pioneering paper by Hans Asperger, "'Autistic Psychopathy' in Childhood." in Frith, Uta, 1991.)

of emotions that cannot be experienced by autistic persons, where deficit in empathy ranks high on the list.[300] Moreover, the recent studies in neuropsychology suggest that both psychopaths and individuals with autism may have similar genetic anomalies at the anatomical level – both conditions have been linked to impairments in amygdala functioning.[301] One would naturally expect that the resulting thoroughgoing emotive disorder would have the most damaging consequences for the capacities of persons with autism to lead moral lives. Yet the available evidence does not always support this expectation.

As Oliver Sacks observed in his well-known account of the time spent with an exceptional scientist and also a diagnosed autistic person Temple Grandin, the lack of the emotive resources that are available to most people does not yet preclude one from developing a mature moral outlook. Temple testifies that from early schooldays she was bewildered by the classic love story of Romeo and Juliet and the story of Hamlet, and Sacks attributes this to her characteristic "failure to emphasize with the characters, to follow the intricate play of motive and intention."[302] He further adds that a number of common emotions, such as diffidence or embarrassment, were quite unknown to Temple. She was wholly lacking in self-consciousness, and even though she could apparently sympathize with the sufferings of animals, she "lacked empathy for people's states of mind and perspectives."[303] As a result, Temple had enormous difficulties understanding the non-verbal, contextual part of the human communication process (e.g., irony or sarcasm), and was truly

---

[300] E.g., Frith, Uta., 1999.

[301] Blair, R. J. 2006; Baron-Cohen et al., 2000.

[302] Sacks, O., 1995, p. 259.

[303] *Ibid*., p. 263, 265; 269.

baffled (feeling herself like "anthropologist from Mars") by that aspect of human life that involved intimate interactions between two people of the opposite sex.

Temple's autism had also a profound effect on her ability to enjoy the aesthetic aspect of the world around her, and, indeed, on her ability to *understand* the basic aesthetic concepts. Sacks records the following incident during his time with Temple:

> We pulled off the road and gazed toward the Rockies – snowcapped, outlined against the horizon, luminously clear even though they were nearly a hundred miles away. I asked Temple if she did not feel a sense of their sublimity. "They're pretty, yes. Sublime, I don't know." When I pressed her, she said that she was puzzled by such words and had spent much time with a dictionary, trying to understand them. She had looked up "sublime," "mysterious," "numinous," and "awe," but they all seemed to be defined in terms of one another.[304]

These various emotive limitations, however, were fully compensated in Temple's moral life by her outstanding abilities to memorize a great variety of situations with a number of characteristic features and apply the same rules of conduct in similar circumstances. Similar cognitive strategies were employed by Temple to decode human emotions – something she was totally unable to recognize intuitively,[305] the way most people do. As a result, Sacks observes, there was hardly anything lacking in her moral stature. Indeed, she conveyed an impression of having a deeper and more profound sense of moral duty than most other people:

---

[304] *Ibid*., 293.

[305] Sacks mentions a report by Dr. Hamelin, of London, in which he told a story about an intelligent autistic girl of twelve, who had come to Dr. Hamelin complaining of the 'funny noises' that her friend Jeanie was making in another room. Upon investigating the matter, it was found that Jeanie was crying bitterly. The significance of weeping (as an expression of sadness or distress) had fully escaped the understanding of that autistic girl – she had merely registered it as something physical – 'funny noise.' (*Ibid*., p. 269). This nicely illustrates the difficulty that autistic individuals experience in understanding the affective side of life. But, as Temple's case shows, this difficulty is yet surmountable.

> Temple is an intensely moral creature. She has passionate sense of right and wrong, for example, in regard to the treatment of animals; and law, for her, is clearly not just the law of the land but, in some deeper sense, a divine or cosmic law, whose violation can have disastrous effects – seeming breakdowns in the course of nature itself.[306]

The case of Temple Grandin is not exceptional. Cathy Grant with colleagues conducted a series of studies of autistic children with respect to their ability to make correct moral judgments, and (contrary to expectations) found that children with autism were as successful as non-autistic children in identifying the morally relevant features of the situation and hence making accurate distinctions between moral and merely conventional rules.[307] Jeanette Kennett likewise reports that "many autistic people display moral concerns, moral feeling and a sense of duty or conscience," whereas "according to all standard descriptions of autism these concerns cannot be based on empathic identification with the concerns and feelings of others."[308] This evidence, in turn, raises an important issue for most versions of the emotive deficiency theory: "If empathy is crucial to the development and exercise of moral agency then why is the autistic person not worse off, morally speaking, than the psychopath?"[309]

Amoralism of a psychopath, as this evidence suggests, is not then *necessarily* due to the deficits in empathy, since, as the example of autism shows, it is at least possible (though, admittedly, much harder) to acquire a level of moral competence utilizing other (presumably, purely cognitive) resources. *How* exactly is it possible is a different question, but Kennett, for

---

[306] *Ibid*., p. 296.

[307] Grant, 2005, pp. 324-5.

[308] Kennett, J., 2002, p. 349. See also Frith (1989) for similar findings.

[309] *Ibid*.

instance, suggests that "capable individuals with autism are likely to take a rather Kantian approach to moral thinking,"[310] i.e., deploying moral concepts of a more dispassionate variety, and developing and rigidly applying the same rules of conduct to similar situations, which can be successfully identified by reasoning alone. Kennett's account corroborates Oliver Sack's impression of how autistic Temple Gardin manages to navigate successfully the puzzling world of morality.

3.3 Empathy and Selective Amoralism

The suggestion that we may need to look beyond the affective aspect of moral performance is further supported by the strong intuition that the inherent *inability* to experience empathy (or any other morally relevant emotion, such as guilt or remorse) cannot be plausibly considered even as a *possible* cause of the immoral behavior in cases of selective amoralists. As we defined this class of amoralists in Chapter Three, the selective amoralists tend to limit the application of moral reasons to a particular group or class of people, but can be quite 'moral' within the limits of that group. 'Being moral' here implies (among other things) that a selective amoralist may experience genuine sympathy with the sufferings of the members of his favorite exclusive group (e.g., family, ethnic group, race, gang, etc.) and on occasions be motivated in the morally appropriate way to alleviate the distress of the 'insiders', thus demonstrating that this important emotive capacity is at least not beyond his reach. As Hanna Arendt reports on the case of Eichmann:

> Half a dozen psychiatrists had certified him [Eichmann] as "normal." One had found that his whole psychological outlook, his attitude toward his wife and children, mother and father, brothers, sisters, and friends, was 'not only normal but most desirable.' […]Eichmann had a conscience, and his conscience functioned in the expected way for about four weeks. […] During those weeks

[310] *Ibid*., p. 350.

> when his conscience functioned normally, it did its work within rather odd limits. His conscience rebelled not at the idea of murder but at the idea of German Jews being murdered.[311]

The apparent problem with Eichmann was *not* his genetic inability to sympathize with the pain of others (as he presumable did when it concerned those close to him), but rather a spectacular failure to extend the scope of this 'moral motive' to the rest of the human race. The specific cause of this failure is what we should inquire after, but it seems more plausible that it lies in the sphere of reasoning about the morally relevant facts and the priority of ultimate values.

Still, not everybody agrees. Arne Vetlesen, whom we have seen earlier presenting a version of the emotive deficiency theory couched in terms of lacking empathy, wants to extend his central insight even to the cases of selective amoralism. Vetlesen refers to one's inability to empathize as 'psychic numbness' – an anomalous suspension of sensitivity to the suffering and pain of others. He argues that this peculiar emotive failing need not be indiscriminate but may be (indeed, typically *is*) selectively applied:

> As for the numb person, the type of indifference he or she shows is neither accidental nor transitory; more psychically profound, numbness signifies a loss of the emotional capacity to develop empathy with alter. But this loss need not be definitive, and numbness is typically highly *selective*, as in the case of the Nazi doctors, where emotional unaffectedness is confined to Jews and other so-called subhumans. Selective numbness is thus organized numbness; it is not original in the subject but brought about in him by the impact of exterior forces. By contrast, indiscriminate or all-around indifference toward other human beings may reveal a more generic psychopathology.[312]

---

[311] Arendt, H. 1963, pp. 22; 90.
[312] Vetlesen, 1994, p. 211.

Vetlesen suggests here that there are two ways in which one can be emotionally numb. First, the defect may be 'generic', which must mean something like the innate inability to experience empathy.[313] Secondly, the emotive deficiency in question might be a case of 'organized numbness', i.e., an *acquired* condition brought about by external forces (e.g., brainwashing), which was 'not original in the subject'. He further makes a point that all cases of *selective* numbness are cases of acquired ('organized') numbness. This numbness, he argues, is precisely what prevents the amoralists from perceiving the intrinsic worth of other people, and thus making a correct judgment about their moral status.

There are several things we can say by way of criticism of Vetlesen's position. It appears that in most cases of selective amoralism (e.g., Eichmann or that of Nazi doctors that he mentions) one's emotive attitude toward the victim is parasitic on one's prior judgment about the victim's status, not vice versa. A Nazi doctor in Auschwitz performing brutal medical experiments on Jewish children has already *judged* that these children are subhuman, and thus do not deserve the same level of concern as his own children, or any other German children. His lack of compassion or empathy toward his victims is a *consequence* of his (mistaken) judgment; and it is not the case that his inability to judge correctly (or to perceive these children as fully human) is a result of his inability to empathize (as Vetlesen suggests[314]).[315]

---

[313] Although Vetlesen never raises this question, but if empathy is indispensable for moral performance, and if come people are born without it (e.g., psychopaths), it is not clear in what sense we can hold them morally responsible for their actions.

[314] "For subjects to arrive at the phenomenon of suffering about which they are to exercise judgment, their emotional faculty (i.e., empathy) as well as their cognitive one (i.e., representative thinking) is inescapably called upon – or else the phenomenon of 'suffering' would not be constituted as an object of *moral* judgment." (p. 158). Vetlesen seems to be saying here that one *cannot* make a moral judgment about suffering (indeed, detect the case of suffering properly) *unless* one empathizes with the victim.

[315] As we have seen, this dependency of emotive state on cognition is present in all instances of non-basic emotions.

This point can be further highlighted by imagining a Nazi doctor who is willing to question his previous judgment, and as a result of further investigation, discovers that there are in fact no morally relevant differences between the Jewish and any other children. Once he realizes this, we can expect that his moral concern will be immediately extended toward the previously excluded group of people, and his attitude toward his job will be drastically changed. We can expect an even more radical change in attitude once he discovers that the child in his concentration camp clinic is his own. Hence, the original problem with an 'unenlightened' doctor was not in his emotive impairments, but rather in his failure to consider all the relevant facts, to question the official dogma, as well as his unwillingness to honestly examine his own beliefs about the matter. The deeper cause of *this* cognitive failure is still something to be determined, but it appears to be largely independent from one's innate predisposition to empathize, Vetlesen notwithstanding. Indeed, the results of the studies cited above find no substantive impairment in moral judgments or judgments about the morally relevant facts among the autistic agents. We would expect an autistic individual with impaired emotive capacities (including the capacity for empathy) to be capable of the relevant inferences required for determining the moral status of children regardless of the children's race or ethnic origin – i.e, precisely the kind of inference that the non-autistic, 'normal' Nazi doctors failed to make.

## 4. Summary and Conclusions

At this point we may summarize the above argumentation and draw several preliminary conclusions. As we have seen in Chapter Three, some extreme amoralists can be plausibly described as being deficient with respect to their emotive capacities. Cleckley, for

example, traces the roots of the anti-social behavior of the psychopath named Joe to the fact that he "does not experience real and serious emotions."[316] Similar descriptions of psychopaths, emphasizing their emotive abnormality, were given by other researchers as well, and there is no reason to doubt their account.[317]

Having healthy basic emotive capacities, furthermore, appears to be indispensable for the seemingly crucial (for moral behavior) ability to experience empathy, which was defined as a partial emotional identification with another person's condition. Yet several considerations, summarized below, prevent us from taking this missing emotional component as the *sole* source of a person's amoralism.

It was argued that the operation of empathy, as (partly) a non-cognitive response to another's emotive state, requires at least a minimal level of intellectual competence. It requires that the person should be able to recognize another's distress, and comprehend the formal criteria of the application of the morally relevant concepts (e.g., pain). Hence, it seems mistaken to emphasize exclusively the non-cognitive aspect of moral behavior and motivation at the expense of the prerequisite cognitive capacities of a moral agent.[318] It further follows that the immoral behavior of a psychopath or a selective amoralist is not *necessarily* the result of the emotive deficiency alone (even if such deficiency does exist).

---

[316] Cleckley, 1982, p. 70.

[317] This does not necessarily mean that the emotional deficiency of amoralists is due to genetic abnormalities. Gary Watson cites interesting facts about the early years of the ruthless killer Robert Harris with all symptoms of severe psychopathy, who was sentenced to death for several cold-blooded murders and was hated even by his inmates in the death-row prison. His sisters testify that as a child "he was the most sensitive of all of us. When he was 10 and we all saw "Bambi," he cried and cried when Bambi's mother was shot. Everything was pretty to him as a child; he loved animals. But all that changed; it all changed so much." (1987, p. 273). This seems to show that the emotive capacities which play an important role in moral responses (even if not a crucial one) were quite intact in him up to a certain point in his development. As Pizarro (2000) has argued, cases of voluntary 'empathic suppression' are not uncommon, and, with continuous effort, a criminal can regulate the 'natural' empathic arousal in order to escape the subjective consequences of his actions.

[318] This minimal level of intelligence might be part of what it means to be a *moral* agent.

This is not to deny that the emotional deficit may be *part* of the explanatory story in certain cases of extreme amoralism. Indeed, well-developed moral emotions (e.g., empathy or sympathetic concern) may often enhance moral performance, and facilitate the recognition of the morally relevant features of a situation. Yet there is no reason to think that a reduced or even non-existent capacity for empathy forever closes for a person the door to the 'world of morality.' The results of studies of autistic individuals show that, in the absence of healthy emotive apparatus, one can successfully utilize various cognitive capabilities to achieve the comparable level of moral maturity and keenness. This, in turn, undermines the common view that the amoral state of psychopaths could be fully explained by their inability to experience certain emotions.

A similar conclusion was reached in cases of selective amoralism. Contrary to Vetlesen's claim that the lack of empathy *determines* one's judgment as to who merits moral consideration and who is excluded from the group, it was argued that the operation of empathy is parasitic on one's prior judgments about the moral status and nature of other people. My argument was further meant to counter the over-inflated role of the emotive factors in normal cognitive operations. The emotive deficiency need not critically affect one's reasoning abilities *per se*.

Admittedly, these results are still inconclusive as to the primary cause of amoralism. If a psychopath is similar to an autistic individual with regard to his reduced (or non-existent) emotive capacities, and this peculiar limitation does not yet preclude an autistic person from acting morally (i.e., in a systematic way), then why is it that a psychopath is so much worse by comparison, morally speaking? This, however, might be a question badly posed. Apparently, not *all* autistic individuals evolve into moral characters, and those who do *not,*

are correctly labeled 'psychopaths'.[319] Likewise, a person with psychopathic tendencies who alters his anti-social behavior (or else, never lets his tendencies control his actions in the first place) is not called a psychopath anymore. Indeed, there are many people who are not very emotional or, in Kant's words, are "by temperament cold and indifferent to the sufferings of others,"[320] but who nonetheless follow their duty with admirable consistency (those rare "good Kantians", as we may call them[321]). Elsewhere, Kant even insists that "virtue *necessarily* presupposes apathy."[322] Thus, the more appropriate question is the following: what is the essential difference between a "good Kantian" and an amoralist – if (*ex hypothesis)* there is no significant difference in their respective emotive conditions? In search for an answer to this question we should now turn to the way the amoralists think and the way they *judge* rather than feel.

[319] See for instance Hans Asperger, *'Autistic Psychopathy' in Childhood*. (in Frith, Uta, 1991).

[320] Kant, I., (1785), 2001, p. 9.

[321] A good historical example of a dispassionate 'Kantian saint' is Maria von Herbert, a contemporary of Kant and a devout follower of Kant's moral philosophy. Her remarkable story is discussed in Langton, Rae, "Duty and Desolation." *Philosophy*, 67 (1992).

[322] Kant, I. (1795), 1964, (*Doctrine of Virtue*, 407). [my emphasis]

# CHAPTER V

## THE COGNITIVE ASPECT OF MORAL AGENCY

> Every wicked man is ignorant of what he ought to do and what he ought to abstain from, and it is by reason of error of this kind that men become unjust and in general bad.
>
> Aristotle, *Nicomachean Ethics*, 1110b27-30.

### 1. Preliminary Remarks

Since the notion of cognitive failure is rather ambiguous, it will be necessary to seek greater precision by identifying the various senses in which a person may be said to exhibit cognitive deficiency of some sort and of various degree. We are primarily interested here in those cases where a failure or inability to reason *properly* (whatever this involves) may adversely affect a person's moral, evaluative judgments (about right or wrong, good or evil) and thus affect his behavior based on those (mistaken) judgments. More precisely, we are interested whether the amoralists exhibit any cognitive deficit with respect to moral practical reasoning which might explain their persistent amoralism. At first glance, the cognitive mistake with moral consequences may happen under at least one of the following conditions:

**(a)** A person may suffer a thoroughgoing failure of intelligence (e.g., a mentally retarded individual, or someone suffering from schizophrenia or other mental disorder of the psychotic group), which prevents him from adequately assessing the morally relevant facts or even from being aware of existence of these facts at all. Such an individual, as it is commonly agreed, would be exempt from the usual requirements of morality as well as from bearing the consequences for his (objectively) immoral acts. Since it has been shown (Chapter Three) that there is no reason to believe that an amoralist (including the psychopath) suffers from a mental impairment in this special medical sense, I shall not discuss these cases any further.

**(b)** A person may exhibit none of the mental disorders mentioned above, and yet have a mind which is not sufficiently developed to allow correct identification of the morally relevant facts or the grasp of even the simplest moral inference, which in turn, might lead to anti-social, destructive behavior. Such is the case of very young children who are not yet competent participants in the world of morality,[323] and it is likewise a common belief that the notion of moral responsibility does not apply here either. Once again, I have suggested earlier (against R. Milo) that children should not be classified under the category of amoralists (they are rather pre-moral),[324] and that in all cases considered in the present work the amoralists are mature adults. Hence, this case of cognitive deficiency will be irrelevant for our discussion as well.

**(c)** A person might have a sound mind in the medical sense, and be of sufficient age to have his mental faculties fully developed, and yet make a mistaken conclusion about any of the following things: the status and the content of the moral principles, the (actual)

[323] Of course, we may have moral duties *toward* children, and in this sense they are part of the moral domain.

[324] My main reason here has to do with the strong negative connotations of the term 'amoralist.'

consequences of his actions, the existence of some morally relevant fact (e.g., the presence of distress in another person), the soundness of a practical syllogism that leads to his particular evaluative judgment (or directly to action), or else, the relative priority between the several conflicting moral values or principles. In any of these cases, it appears, the mistake may lead to an action that openly defies the requirements of morality. And when the mistake is fundamental enough in the sense that it indeed results in a systematic immoral conduct (not just occasional moral lapses) or makes a person wholly insensitive to reasons derived from moral consideration, we may then point to it as the primary explanatory cause of the amoralist's condition.

It is perhaps overly presumptions to expect that all amoralists commit one *specific* kind of mistake in their moral considerations, but if any such mistake is (typically) committed, and if it causally contributes to actions, we should be able to describe it in terms general enough to account for most, if not all, interesting cases.

## 2. The 'Cognitive Failure' Theory

It is not uncommon to hear a person justifying his or her moral disapproval of someone's destructive behavior toward others by posing the following hypothetical question: "How would you like if you were in the place of the person you are hurting now?" This frequent strategy is indicative of an existing intuition that acting morally requires the ability to assume the perspective of other agents, and is perhaps pointing to an even stronger belief that a successful role-taking is *sufficient* for making a morally praiseworthy choice. It suggests that achieving (at least) the second stage of universalization (on Mackie's scale[325]) provides or

[325] See Chapter 2.

should provide a sufficient *motivating* reason for acting morally. Since the kind of role-taking that is required by the second stage universalization is clearly a cognitive skill that normally matures with the development of general intelligence, the inability to assume the perspective of others and, as a consequence, failure to act in accordance with the requirements of customary morality may point to a specific cognitive deficiency on the part of the immoral or amoral agent.

In her attempt to understand the roots of Eichmann's amoral nature, Hannah Arendt observes that "a more specific flaw in Eichmann's character was his almost total inability to look at anything from the other fellow's point of view."[326] Even though Arendt apparently does not suggest this to be the only flaw in his character, she clearly believes that a certain cognitive deficiency[327] on Eichmann's side can be at least partially explanatory of his criminal behavior during the Holocaust:

> The longer one listened to Eichmann, the more obvious it became that his inability to speak was connected with an inability to *think*, namely, to think from the standpoint of somebody else. No communication was possible with him, not because he lied but because he was surrounded by the most reliable of all safeguards against the words and the presence of others, and hence against reality as such.[328]

The most surprising feature of Eichmann's story, for Arendt, was the impossibility to see Eichmann as a truly wicked or evil person, despite the fact that his actions resulted in unprecedented amount of suffering. It appears, she suggests, that Eichmann was not wicked, but rather *thoughtless* – "he *merely*… never realized what he was doing." This

---

[326] Arendt, H. 1963, p. 43.

[327] Or deficiency of *thought*, as Arendt prefers to call it.

[328] *Ibid*., p. 44.

thoughtlessness[329] that prevented him from judging correctly and led him into doing evil was the result of his inability of representing others in his own mind, which includes his inability to *universalize* the value judgments in the morally required way. This is precisely the sense in which Eichmann, a selective amoralist, was cognitively deficient, according to Arendt's analysis.

A psychologist Lawrence Kohlberg echoes to some degree Arendt's neo-Kantian intuitions when he argues that the capabilities for complex perspective taking (i.e., cognitively taking the perspective of another), and for understanding abstract, morally relevant conceptions underlie advances in moral reasoning and in the quality of one's pro-social (moral) behavior. According to Kohlberg, there exists a direct relation between advances in general intelligence and moral progress in most normal individuals – e.g., children are expected to develop higher-level moral reasoning as they become capable of cognitive skills required for understanding such reasoning. Correspondingly, a failure to advance to a higher moral stage might be accounted by a failure in cognitive progress, including the failure to grasp the meaning of moral terms.[330]

In a similar vein, Stanton Samenow in his comprehensive work "Inside the Criminal Mind" probes into the inner world of a compulsive criminal (whom he explicitly identifies with a clinical psychopath)[331] and concludes that what fundamentally separates a psychopathic criminal personality from the rest of us is the difference in a way of *thinking*:

---

[329] *Ibid*., p. 287.

[330] Kohlberg, L. 1973; 1976. Kohlberg postulates the six stages of moral reasoning, where each subsequent stage marks a progress in moral development which directly depends on the level of intellectual maturity of an individual. E.g., only a select group of individuals who had formal training in ethics would be capable of assuming a (deontological) moral point of view as described by the highest (sixth) stage.

[331] "Although diagnosticians may make distinctions between the psychopath and criminal, for all ostensible purposes, one differs hardly at all from the other." (Samenow, S. 1984, p. 181.)

> We are as we think. It is impossible to help a person give up crime and live responsibly without helping him to change what is most basic – his thinking. Criminals have been rewarded, punished, manipulated, probed for unconscious dynamics, and taught to read, work, and socialize, but they have not been helped to learn brand-new thinking patterns in order to change their way of life. It is critical that all of us know who the criminal is, that we realize that he thinks and acts differently from the rest of us.[332]

In all these cases the emphasis is made on some cognitive shortcoming (or at least, a cognitive peculiarity) on the part of the amoralists, and this failure is variously specified as either inability to assume the point of view of another, as the lack of mastery of the moral concepts, or else as some deeper cognitive defect that involves distortions on the most fundamental levels of thinking and judging. It will be convenient to refer to this very broadly defined position as the Cognitive Failure Theory. This position can be further elucidated by focusing on the key concept of 'rationality.'

2.1 Acting Rationally: Let Me Count the Ways…

'Rationality' is one of the many notoriously ambiguous philosophical terms that admit of several diverse and even mutually exclusive interpretations. If there is any hope of understanding the contention that amoralists are 'in some sense' deficient in rationality (as the cognitive failure theory entails), we must clarify this key notion first.

To begin with, we may observe that rationality (as well as its contrary, irrationality) is commonly attributed both to actions and to a process or result of purely theoretical reasoning about facts. The rationality of actions, however, seems to be dependent on the practical deliberation that leads to actions. The traditional distinction between practical and theoretical

---

[332] *Ibid*., p. 257. I believe though, that Samenow should have made a stronger statement here. As I have argued in Chapter Two, Section 3, the amoralist is not just *different* in the way he thinks and acts, but we should also be able to say that there is something *wrong* with the amoralist. If Samenow is correct, and the amoralist's thinking is the primary explanatory cause of his condition, we must then show that his thinking is indeed *mistaken* or *deficient* in some sense.

reasoning emphasizes the difference in goals. Whereas theoretical reasoning aims at understanding, its practical counterpart seeks to choose between any number of alternative actions, i.e., its goal is connected with decision and behavior. Success of theoretical reasoning is often described in terms of presenting strong reasons for *believing* a certain proposition, while practical reasoning, if successful, should provide good reasons for *acting* in a certain way. Given the difference in goals, the criteria of evaluation for the two kinds of reasoning will obviously differ.

It is perhaps safe to assert that, historically, disproportionate attention has been paid to the standards of evaluation and the objective criteria of assessment of theoretical arguments, but much less has been written on the logic of practical reasoning. We can quickly recall some of the common fallacies, formal or informal, which would undermine the cogency of an argument, and, in turn, make one's acceptance of the conclusion of this argument unwarranted or (in some cases) irrational. It is much harder to specify the conditions under which an action would be equally irrational, or done without good reason. Are there rational constraints on the *goal* of an action, or is rationality of an action simply a function of its instrumental value in achieving any goal the agent cares to adopt? The available responses to this question will probably show much less consensus as compared to the question what makes an argument valid or inductively strong.

On the surface, the most obvious sense in which a person may act contrary to reason, and thus exhibit a level of irrationality, is a case when his voluntary actions are detrimental to his own *goal*. Consider a person who wants to regain his health *and* knows that a pill *X* will cure his ailment, (and he has no ideological or religious scruples against modern medicine), but nonetheless refuses to take the pill. His behavior may rightfully be describes as

*irrational*, and we may further inquire what caused his practical reasoning to fail in this case (e.g., weakness of will).

The ascription of irrationality becomes less straightforward once we complicate our example a little. Consider a person who wants to be cured, and who has the efficacious pill *X* in his possession, but *does not* know of the pill's powers. As a result, he refuses to take the pill. Apparently, from the objective perspective, his refusal to take the pill was detrimental to his own goal. But did he act irrationally in this case? We would normally say that the cause of his imprudent behavior was *ignorance* rather than irrationality, since this last term tends to have stronger negative overtones. Ignorance, nonetheless, is a kind of cognitive deficiency; an ignorant person falls short from a certain level of intellectual excellence. We may agree that not all cases of ignorance are blameworthy, but, as a popular idea of the perfectly rational being suggests, all cases of ignorance manifest a degree of intellectual inadequacy.

In both examples the rationality of action (or inaction) was evaluated against the preset end (namely, health). In both cases, the end was perceived as something *good* for the agent. An irrational person, in this sense, is a person who fails to take the most adequate means for achieving his end. Irrationality is most noticeable when the person *knows* which steps must be taken, is capable of taking these steps, but fails to act accordingly. In the second example, the relevant knowledge is absent, but this ignorance can still be construed as a failure of rationality in some weaker sense. The same reasoning would apply to a case when a person is ignorant of what is *really* good for him, rather than of the means to achieve this good.

Finally, we can imagine a person who consciously sets a goal that is detrimental to his own good (however this good is specified). For argument's sake, we may assume that it is

possible that a person desires, for instance, to ruin his health for the sake of ruining his health and for no further reason.[333] He may also agree that being healthy is in all senses *better* than being sick, and yet stick to his goal. Furthermore, he also knows that taking some other pill *Y* is the most direct route to his self-destructive goal. And so he takes the pill. Is there anything irrational about this person's behavior?

Apparently, there is no failure of *instrumental* rationality here: he takes the best means to achieve the end. His action of taking the pill is perfectly rational insofar as it promotes his goal. If anything, it is the rationality of the goal itself that might be questioned. But against what standard can we determine the rational status of one's goals? What could serve as the non-arbitrary criterion of rationality in this case?

Answering these questions is crucial for the further discussion of the amoralists' alleged failure of rationality. Arguing that one's goals are subject to rational evaluation from some objective standpoint amounts to saying that there is one point of view of reason *sans phrase,* and that this unique standpoint, once fully specified, is the ultimate reference for all normative claims about the ends of action. If, on the other hand, the universal test of rationality cannot be applied to the chosen end, but only to the means to this end, then we can properly speak only of what is rational to do from a particular point of view, i.e., the point of view of a certain system of values. But the original act of *adoption* of this value system itself will remain beyond rational assessment. Furthermore, if the notion of rationality is exhausted by its instrumental variety, the same action may be both rational and irrational, depending on

[333] For the argument that tries to establish this possibility see Michael Stocker, "Desiring the Bad" (1979). Dostoyevsky in his "Notes from Underground" shows a literary character who epitomizes this kind of self-destructive consciousness: "I am a sick man… I am an angry man. I am an unattractive man. I think there is something wrong with my liver. I am not having any treatment for it, and never have had, although I have a great respect for medicine and for doctors. […] No, I refuse treatment out of spite. […] I know better than anybody that I am harming nobody but myself. All the same, if I don't have treatment, it is out of spite. Is my liver out of order? – let it get worse!" (1988, p. 15).

the assumed frame of reference. Cleverly cheating on one's taxes may be rational from the point of view of prudence (i.e., where the goal is maximizing one's own interest), and irrational from the moral point of view (i.e., where the goal is, for instance, maximizing the welfare of the greatest amount of people). Since the amoralist cares little about moral values, the charge that his anti-social actions are irrational from the point of view of morality, will be trivially true (i.e., in a hypothetical world where, say, overall happiness is one of the amoralist's goals), but hardly relevant to his own actual condition.

In what follows I shall address these questions in the following order. I shall first inquire whether amoralists are (typically) impaired in their ability to use instrumental (means-end) reasoning, as compared with normal subjects. It will be argued that the available evidence shows no such impairment. I shall then raise the issue whether instrumental rationality is the only kind of rationality, or whether we can give a *substantive* account of reasons for action which would apply to all humans regardless of their chosen goals or preferences. An argument will be made to support an affirmative answer to this last question, even admitting that there can be no decisive proof of this claim. Lastly, I shall ask whether anti-social, immoral behavior of an amoralist goes against these universal reasons for action, and can thus be properly described as irrational.

2.2 Means-End Rationality and General Intelligence

As it was stated earlier, instrumental, or means-end rationality, involves the ability to determine the most efficient means for achieving one's end. As such, it clearly requires from the agent the possession of a degree of general intelligence as well as competence in basic theoretical reasoning skills. "Intelligence" generally consists of different abilities, such as the ability to reason logically, solve problems, think abstractly, learn and understand new

material, adapt to novel situations quickly, grasp complex relationships, and profit from past experience. It is a common place that some people are much more successful at this kind of reasoning than others. But the question we want to consider in this section is whether there is any evidence that shows psychopaths, hardened criminals and other amoralists as being *in general* less competent at these cognitive tasks. In other words, are there reasons to believe amoralists to be inherently *less* intelligent than the rest of the human population?[334]

In fact, there exists overwhelming empirical evidence that a great number of individuals suffering[335] either from the antisocial personality disorder or from other forms of selective amoralism exhibit high levels of general intelligence. Indeed, in some cases the level of intelligence exhibited by a psychopath is significantly higher than that of an average person of a comparable age and social status,[336] and the image of a 'bright sociopath' is widely exploited in popular fiction and film.[337] Some evidence to the full intellectual competence (or even brilliance) of many amoralists has already been cited in my earlier

---

[334] I am not interested here with the question whether *ordinary* criminals are less intelligent than the law-abiding citizens. If IQ tests are indicative of intelligence, there is some evidence that shows most criminals scoring slightly less on these tests than the majority of the population. The significance of this fact is questionable, since the lower scores can often be explained by lower socioeconomic conditions of the delinquents, and the strictly causal relation between crime and intelligence has not been established.

[335] The use of the word 'suffering' in this context is idiomatic rather than literal. As Martha Stout rightly observes, "sociopathy stands alone as a 'disease' that causes no *dis-ease* for the person who has it, no subjective discomfort. Sociopaths are often quite satisfied with their lives." (2005, p. 12)

[336] Admittedly, there exists some evidence to the contrary. Schelling and Rosen (1968) report that psychopaths perform abnormally on some cognitive tasks, especially involving processes that are associated with frontal lobe functioning. For example, they tend to make more errors than normal subjects when asked to complete Porte's mazes of increasing difficulty. However, a more recent study by Patricia Sucker *et a*l. (1983) puts these results in doubt. After studying the responses of the 44 incarcerated psychopaths as compared to those of normal subjects, she found that "compared to normal controls, psychopaths showed no greater tendencies to persist in incorrect responding, disregard cues suggesting the need to modify behaviors, or proceed impulsively in problem solving." (p. 275). Overall, she concludes, it appears that "deficits in cognitive functioning are not *necessarily* characteristic of psychopaths across situations" (*Ibid*.) (my emphasis).

[337] E.g., Stanley Kubrick's film "A Clockwork Orange" (1971) based on Anthony Burgess' novel.

discussion of the specific cases of amoralism in Chapter Three. But two further examples should suffice to illustrate this point.

A quite dramatic case that tends to undermine the claim about the essential connection between morality and general intelligence is described by Daniel Tranel in his discussion of a condition known as "acquired sociopathy."[338] This condition develops after lesions to the frontal lobe of the brain in individuals whose personalities were entirely normal prior to the operation. Patient, known in literature by his initials EVR, was a successful hard-working professional, happily married and the father of two well-adjusted children. He led an impeccable social life, was considered a leader in his community and a role model for his younger siblings. At age 35 EVR underwent an operation to remove a brain tumor, which involved bilateral excision of orbital frontal cortices. Even though the operation was a success, and the threat of brain cancer was eliminated, the effect the operation had on EVR's personality stunned everyone who knew him. Entering several disastrous business ventures, divorcing his first wife, remarrying and divorcing again soon after, being unable to plan his future activities or to hold any job due to continuous irresponsibility on his side, and ultimately ending up in a sheltered environment – such biographical entries are indicative of a typical path of a clinical sociopath. Tranel further observes that EVR's moral sense has been affected as well and in a peculiar way:

> His ability to decide which persons are good and which are bad, which was previously superb, is now virtually absent. Similarly, he lacks a sense of what is socially appropriate, although it is obvious from his premorbid life and achievement that he once had a keen sense of social appropriateness.[339]

---

[338] Tranel, D. 1994. This condition is correlated with the more usual cases of the *developmental* sociopathy.

[339] *Ibid*., p. 287.

The symptoms described allowed the psychiatrists to diagnose him with the antisocial personality disorder that was acquired (most likely) as a result of a surgical brain damage. But what is most significant in this context is the fact that the brain lesion that resulted in such radical personality changes left the patient's intellectual abilities apparently intact:

> The change is made the more dramatic by the fact that his basic intellectual and cognitive skills were unaltered by the operation. He continues to perform in the superior range on conventional tests of intellect […], it is abundantly clear that he can distinguish with great subtlety among highly ambiguous concepts, use deduction and induction fluently […]. His speech and language processing are fully preserved, and conventional learning and memory capacities are also intact.[340]

How can we possibly account for cases like this? After all, EVR seems to retain not only his general intelligence, but, more specifically, there is no reason to think he had suddenly lost his mastery of the psychological concepts or lost the capacity to represent the mental states of other people (role-taking capacity). Yet, both the quality of his moral judgment ("deciding which persons are good and which are bad"), his ability to cope effectively with social situations, and his external behavior have deteriorated dramatically.[341]

The startling case of 'moral degeneracy'[342] without any visible deterioration in reasoning ability can be observed on the example of the whole communities as well. Colin Turnbull's well-known account of the Ik people's moral deterioration in Northern Uganda is a case in point.[343] Turnbull explains that the hunter-gatherer tribe of Ik started to get into difficulties at the end of the World War Two as the new nation states in the area "hardened"

---

[340] *Ibid*.

[341] EVR's case is not unique. See Greene (2005) for a very similar account of a patient named Elliot.

[342] A term introduced by Hendry Maudsley in the beginning of 20th century.

[343] Turnbull, 1987.

their frontiers. Previously the Ik had moved freely between several African countries hunting game in the Kidepo valley and following it through the mountainous areas into the Sudan. However, the Kidepo area was declared a Ugandan National Park and they were no longer allowed to hunt there. As a result, they were forced to make a permanent base on the unwelcoming mountainous eastern edge of the Park – an environmental change that affected dramatically both their customary lifestyle and the social relations within the group. Christopher Cherry gives a vivid summary of Turnbull's report on the present *amoral* condition of the Ik tribe:

> A child dumped on the ground is seized and eaten by a leopard. The mother is delighted; for not only does she no longer have to carry the child about and feed it, but it follows that there is likely to be a gorged leopard near by, a sleepy animal which can easily be killed and eaten. An old woman who has been abandoned falls down the mountainside because she is blind, so a crowd gathers to laugh at the spectacle of her distress. A man about to die of gunshot wounds makes the last request for tea. As he feebly raises it to his lips, it is snatched from him by his sister, who runs away delighted. A child develops intestinal obstruction; so his father calls in the neighbors to enjoy the joke of his distended belly.[344]

This drastic weakening of all social ties and the breakdown of all previously honored norms and values was the result of forced changes, Turnbull suggests, in Ik's traditional way of life. Prior to these geographical and ecological changes, the customary morality of the tribe was no different and no less scrupulously observed than that of the other related communities. The collapse of the orderly social existence was swift, deep and seemingly irreversible.[345] Yet there is no evidence that Ik's average level of intelligence suffered a

---

[344] Cherry, 1975, p. 73.

[345] Turnbull himself holds a rather pessimistic view on the possible improvement among the Iks, and suggests that the government should break up the remainder of this community, and resettle small groups of Iks among other tribes.

decline in any way comparable to their moral deterioration. Indeed, the tribe is known by its neighbors as being rather inventive and cunning when it comes to stealing their cattle.

These two examples should suffice here, and we may draw some preliminary conclusions at this point. Referring to 'feeblemindedness' as the main cause of the amoral (as well as immoral) condition was once a common strategy. Indeed, up until 1928, the American government accepted the idea (popularized by H. H. Goddard[346]) that all feebleminded people were potential criminals, and required that all new immigrants take the intelligence tests. Imbecility was seen as the primary cause of ordinary criminal activity, but, by analogy, was extended to cases of extreme amoralism as well. Overall, however, given the available evidence, we may conclude that a popular conception of hardened criminals, psychopaths and other amoralists as being (on average) less intelligent than ordinary people of comparable social position but who recognize the constraints of morality is mistaken. Since general intelligence is what is required for successful means-end reasoning, the amoralists (on average) are as likely to reason and act rationally or irrationally as other people. Persistent instrumental irrationality cannot be seen as a distinctive mark of the amoralists, let alone as the primary cause of their condition. Hence, if failure of cognition nonetheless underlies the amoralist's condition, it must be sought in some other aspect of this multifaceted faculty.

2.3 Amoralism and Moral Incompetence

An alternative, although a related, way in which one might attempt to explain the condition of an amoralist from the perspective of cognitive failure is by construing such a person as being morally *incompetent*, by analogy with cases of incompetence in other areas

---

[346] Goddard, 1915.

of human expertise (e.g., the incompetence of an illiterate person). We must then examine in what sense a person may be said to lack moral competence, leaving aside cases where such apparent incompetence is due to intellectual immaturity (e.g., in young children), 'feeblemindedness', or mental disorder (e.g., in psychotics).

The (alleged) moral incompetence of amoralists may be further attributed to a number of causes. First, (assuming morality is not innate) we may think of a situation where one was never *taught* about moral rightness or wrongness, and thus he lacks the necessary knowledge (e.g., knowledge of moral rules) for making the correct distinctions in real life. The existing reports of children brought up by animals in the jungle would be the most conspicuous cases of the lack of moral education; but, surely, cases of neglect in moral upbringing of children are not uncommon in many dysfunctional families. This kind of moral ignorance would then straightforwardly explain the persistent immoral behavior, and could further serve as a mitigating circumstance in judging the amoralist's anti-social conduct.

However, all cases of amoralism we considered in the present work, are *not* cases of moral ignorance in this sense. Evidence was cited showing psychopaths successfully *invoking* the moral point of view when it serves their interest, and being capable of identifying correctly the instances of conventional moral categories (e.g., right or wrong, a benevolent or cruel action, etc.). In most examples (with the possible exception of aggressive psychopaths and serial killers) the amoralists mentioned were brought up in the usual social settings, being routinely exposed to conventional moral rules and values. As Murphy rightly observes, "psychopaths know, in some sense, what it means to wrong people, to act

immorally…"[347] Hence, it is highly unlikely that reference to ignorance of the content of morality could play an explanatory role in the overwhelming majority of cases of amoralism.

Secondly, (and, admittedly, more exotically) one might claim that he simply *forgot* what rightness and wrongness amounted to, even if he received moral education in the past. The possibility of forgetting the difference between right and wrong rests on a questionable assumption that moral knowledge is exhausted either by a kind of *propositional* knowledge about facts or else a *competence* knowledge that is characteristic of technical skills. On this assumption, one's knowledge of right and wrong may get rusty over time and with only few occasions to practice it in exactly the same way as one's knowledge of Latin may get rusty without sufficient linguistic exercise. I believe we can agree at this point with Gilbert Ryle (even if his behaviorist account of moral knowledge is much less plausible) when he writes:

> Virtues are not skills; they should be classed rather with tastes and preferences, and particularly with educated tastes and cultivated preferences. Taste is educated preference, preference for recognized superiorities. To be able to recognize superiorities is to know the difference between good and bad.[348]

The ability to recognize a morally superior course of action and choosing this course is not merely mastery of a difficult technique or acquisition of certain beliefs (although beliefs play a role here). As Ryle puts it, "there is nothing in particular that the honest man knows, *ex officio*, how to do." Being moral essentially involves having the right *preferences*, even if it is not *all* there is to being moral. Preferences may change over time but there is an incongruity of sorts in saying that one has forgotten what one prefers or what one values. Hence, we may safely conclude with Ryle:

---

[347] Murphy, 1972, p. 287. Cf. also R. Hare's observation that "psychopaths *know* the rules but follow only those they choose to follow, no matter what the repercussions for others." (Hare, 1993, p. 75)

[348] Ryle, 1958, p. 384.

> It is ridiculous to say that one has forgotten the difference between right and wrong. To have been taught the difference is to have been brought to appreciate the difference, and this appreciation is not just a competence to label correctly or just a capacity to do things efficiently.[349]

I shall therefore leave aside this particular interpretation of the cognitive failure involved in amoralism.

A more sophisticated interpretation of the kind of incompetence involved in amoralism refers to a cognitive failing at the level of moral *concepts* rather than moral principles or values. It might be argued that amoralists lack understanding of the basic moral categories, such as rightness or wrongness, or, perhaps, of some of the important morally relevant concepts, such as cruelty, suffering, pain and pleasure. In assessing the plausibility of this alternative, we should first clarify the conditions necessary for grasping some of these concepts.

To begin with, the normative concepts of wrongness and rightness, good and evil are usually defined against either the background of moral rules or moral values. E.g., a morally wrong action is the one that violates the relevant moral principle, and a good or virtuous conduct is the one that promotes a certain moral value. Hence, grasping the meaning of 'rightness' and 'wrongness' depends on knowing the relevant moral prescriptions and prohibitions, or else, being aware of what is valuable *from* the moral point of view (even if one refuses to adopt this perspective himself). It follows that the claim to ignorance of moral concepts will be a variation of the general moral ignorance claim discussed above, and all the same points would apply here as well.

'Cruelty', 'suffering', 'pain', and a number of related terms are admittedly very different from the purely normative categories mentioned above. In addition to having the

[349] *Ibid*., p. 387-88. This idea, of course, can be traced back to Aristotle.

*prima facie* evaluative overtones, these concepts are partially descriptive as well. A number of philosophers make a distinction between the so-called 'thin' moral terms (e.g., 'good', 'right', 'ought'), which say little or nothing specific about the agent, action or the state of affairs that is being so appraised; and the 'thick' moral terms like 'brave', 'generous', 'merciful', 'cruel', which, in addition to being evaluative, contain substantive criteria for the term's application.[350] Thus, these latter 'thick' terms essentially consist of two elements, the normative and the factual or descriptive. We have seen that knowledge of the normative aspect of a term (e.g., whether cruelty is a good or a bad thing) is derived from the knowledge of normative rules and awareness of the relevant values. Grasping the purely descriptive criteria of the term's application, on the other hand, requires nothing more than being a competent member of the language community. This last prerequisite, in turn, is guaranteed by having an unimpaired general intelligence, which we have seen no good reason to deny to most amoralists.

There remains one other worry that needs to be addressed before we can dismiss the moral incompetence interpretation of amoralism. Antony Duff in his discussion of psychopathy describes (and later argues against) a philosophical position which supposes that providing and mastering purely descriptive criteria for the application of the evaluative moral terms is sufficient for acquiring full moral competence.[351] The defenders of this view rest their case on the empiricist assumption that the intellectual grasping of the moral concepts and categories is no different in principle than the process of acquisition of any other kind of concepts and categories. As with other complex cognitive structures, mastering some higher

[350] Cf. Hare, R. M., 1952; Williams, B. 1985, p. 143-5; and McGinn, C., 1997, pp. 92-3.

[351] Duff, A. 1977, pp. 197-99.

moral concepts will require the mastery of a number of other, non-moral concepts first, but the moral learning process, on this view, follows the usual patterns of progressive understanding and involves incorporating the earlier (more basic) stages into the higher (more involved) structures. Thus, Colin McGinn suggests that mastery of folk-psychology constitutes both necessary and sufficient condition for the full moral competence:

> There could be no such thing as a mastery of moral concepts that did not involve a mastery of psychological concepts. But by the same token it is hard to see how a mastery of folk psychology, including the concept of objective existence of distinct persons, could fail to lead naturally to a grasp of moral categories. Once you have a concept of pain, and you have general reflective intelligence, then you are virtually bound to see that it is a bad thing – for others as well as yourself.[352]

This quote is interesting for a number of reasons, but I shall limit the discussion only to the points directly relevant to the theme of the present section. Strictly speaking, McGinn does *not* suggest here (or elsewhere) that all amoralists (or immoralists) necessarily lack the mastery of moral concepts. What he does say, though, is that *if* amoralists are cognitively deficient at the level of moral categories, it *must* be because they failed to acquire a mastery of folk psychological concepts. The moral failure then is supervenient on (and is indicative of) the more fundamental cognitive limitation. Furthermore, a mastery of folk psychology (combined with "general reflective intelligence"), according to McGinn, *guarantees* sufficient competence with respect to moral concepts. Finally, I take it that McGinn's thesis is not meant to be an *a priori*, conceptual truth about the *meaning* of the moral (or psychological) terms, but rather an empirical observation about the natural dependency between the one type of knowledge and the other.

---

[352] McGinn, 1997, p. 56.

I believe we can reasonably question both claims that McGinn makes in this short passage; namely, the claim that mastery of folk psychology is *necessary* for moral competence, and the claim that it is *sufficient*. The last contention is an easier target, since we can readily imagine an intelligent extraterrestrial being or a perfect sadist who is capable of recognizing the reality of other people, understands well enough the notions of belief, intention and desire, agrees with the factual claims of pain ascription, but disagrees with the (usual) evaluative pronouncements about pain, without yet betraying a lack of "general reflective intelligence." It is conceivable for a sadist to recognize that his victim is in pain, and even agree (with McGinn) that being in pain is a *bad thing* for the victim, without yet *caring* in the least about alleviating the victim's suffering. Such a person, we should say, is not yet part of the world of morality, and whatever formal mastery of the moral concepts' usage he has, it is not sufficient for genuine understanding of what these concepts mean for moral agents.

The claim that takes a mastery of folk psychological notions as a *prerequisite* for moral competence is perhaps more plausible. After all, it is not easy to construe a mature moral agent who lacks basic awareness of other person's mental states, such as intention, beliefs, expectations and desires. Awareness of these states of the 'other' seems to be essential both for judging the morally relevant situation correctly and acting in accordance with one's judgment. The oft repeated claim that amoralists are unable to represent the world from the other person's perspective is part of the same concern. The difficulty experienced at this kind of universalization was observed earlier both in psychopaths and selective amoralists like Adolf Eichmann.

Nonetheless, there is some credible evidence suggesting that, contrary to McGinn's thesis, mastery of psychological concepts might not be *necessary* for moral competence after all. The evidence, once again, comes from the studies in autism. Cathy Grant with a group of scholars has conducted studies of children with autism who have, in addition to emotive deficiencies, an impaired understanding of mental states and of other minds.[353] Prior to experiments, the prediction was made that deficiencies in understanding of mental concepts will adversely affect the ability to make correct moral judgments, distinguish moral from the social (conventional) rules, and identify the morally relevant features of the hypothetical situation. The central finding of the study, however, was that children with autism showed some overall unexpected abilities on the various moral judgment tasks:

> Children with autism were as likely [as non-autistic children] to use motive for the basis of culpability judgments and did so even when the outcome of behavior was negative. Children with autism were also as likely as children in other groups to judge damage to people as more culpable as damage to property. […] Children with autism were able to provide some appropriate justifications for their judgments.[354]

Admittedly, the experiment described here deals primarily with moral *judgments* rather than with moral reactions or moral motives. But we may suppose (on the basis of evidence cited earlier) that given the right circumstances and barring the weakness of will, the correct moral judgment of an autistic child would manifest itself in moral behavior. Hence, it would be correct to describe these autistic individuals as being morally competent *despite* their limited (or even non-existent) grasp of psychological concepts.

---

[353] *Diagnostic and Statistical Manual for Mental Disorders* (DSM IV) describes autistic disorder as a condition involving (among other things) lack of social or emotional reciprocity. Besides, "autistic individuals' awareness of others is markedly impaired, […] they may be oblivious to other children, may have no concept of the needs of others, or may not notice another person's distress." (DSM IV, Article 299.00)

[354] Grant et al., 2005, p. 325.

Evidently, then, the condition of amoralists is unlikely to be linked to the essential failure of instrumental reasoning due to poor general intelligence, or else, to any version of the moral incompetence theory discussed in this section. Admittedly, even if a strict correlation could be established between amoralism and cognitive failure in any of the above senses, it would not yet provide a sufficient reason for taking cognitive deficiency as a *cause* of amoral state, and an additional argument to this effect would be needed. But given the results of the above reasoning, we can safely dismiss the view that impaired reasoning ability is directly responsible for one's failure to assume the moral point of view, or inability to take the moral reasons seriously enough. This, however, does not preclude the possibility that amoralists are *irrational* in some other, non-instrumental (and non-technical) sense of rationality. The following section will explore this option.

## 3. The Substantive Constraints on Rationality of Action.

The results of the preceding sections suggest that the source of (alleged) irrationality of amoralists must be sought outside the spheres of failed instrumental reasoning, moral ignorance or incompetence at the level of moral concepts. But in what other sense can one's practical reasoning (as well as actions based on it) be rationally deficient?

### 3.1 Robert Audi's Account of Rationality of Action

Historically, the accounts of rational intentional action stemming from Aristotle, Hume and Kant have been most influential. Robert Audi, in his latest book, argues that all three philosophers, despite their important differences elsewhere, in their explanations of intentional action accepted what he calls the *correspondence thesis*: "The view that to every intentional action there corresponds at least one practical argument whose premises express

motivation and belief jointly sufficient to explain the action."[355] The truth of the correspondence thesis does not imply that every intentional action *in fact* arises from practical reasoning (clearly, most of our actions are done without much deliberation), but the claim that the relevant practical argument *can* be reconstructed in all cases of intentional actions provides the necessary theoretical background for the possibility of critical assessment of practical reasoning, and eventually for the specification of the conditions of a rational action. If intentional action is equivalent to action for a reason, we may evaluate the action itself through the evaluation of the corresponding practical argument.

On Audi's account, a paradigm example of a practical argument contains the motivational clause in the major premise (e.g., "I want to **φ**"), a belief statement in the minor premise (e.g., "My X-ing would contribute to realizing **φ**"), and the conclusion, consisting either in normative judgment (e.g., "I should X") or the action (X-ing) itself.[356] The next question, then, is what conditions need to be satisfied in order for an action issuing from the corresponding practical argument to be rationally grounded?

In approaching the assessment of practical reasoning Audi considers at least four sorts of criteria: the logical, epistemic, material, and the inferential criteria. The first three criteria are all familiar evaluative practices that apply both to theoretical and practical arguments. The logical criterion governs the logical relation of the premises to the conclusion (e.g., whether the argument is valid or inductively strong), the epistemic criteria concern the kind of justification we can provide to make the premises generally acceptable, and the material condition inquires into the *actual* truth of the premises. The inferential criterion, on the other hand, seems to be relevant specifically to the appraisal of practical reasoning. The inferential

[355] Audi, 2006, p. 28; cf. also p. 103.

[356] *Ibid*., p. 141. Audi's interpretation of Aristotle's position here leaves room for both kinds of conclusion.

criterion concerns the reasoning *process* rather than the logical relation between the premises and conclusion. In effect, it requires that a practical judgment (conclusion of a practical argument) that the arguer holds (typically, a belief) is, as a matter of fact, accepted *because* of the logical force of the corresponding practical argument. The case of *rationalization* illustrates the type of reasoning that fails on this last criterion. We have an instance of rationalization when one holds the practical conclusion on some basis *other* than the premises of the (presented) practical argument. The argument, in other words, does not causally *explain* the acceptance of a belief.[357] In these cases, Audi argues, no matter how high the argument scores on the other three criteria, "the reasoning fails to produce knowledge, or justified belief, of that conclusion."[358]

The analysis of practical reasoning together with the methods of its evaluation has an important bearing on the rationality of an intentional action that is the product of such reasoning. The connection between a sound practical argument and the rationality of the ensuing action is straightforward, according to Audi:

> Ideally, good practical reasoning expresses a valid underlying argument with premises that are true and justifiably believed, and with a conclusion which is both justifiably inferred from them and justifiably held on the basis of them. […] When an agent makes a practical judgment on the basis of practical reasoning that fulfills or approximates this ideal, that judgment and action on the basis of it will tend to be rational.[359]

---

[357] Audi makes a helpful reference to Kant's distinction between acting from duty and acting merely in accordance with it, which nicely illustrates the *inferential* failure (pp. 145-48). A shopkeeper who deals honestly out of selfish motives is acting in accordance with a maxim that *could* have been arrived at by a valid moral argument. But that is presumably *not* why he accepts this maxim. We may say that a maxim is *inferable* from the moral argument, but it was not so *inferred*.

[358] *Ibid*., p.145.

[359] *Ibid*., p. 157.

Of course, this is merely a formal description of a rational action which by itself will not yet help one to distinguish between rational and irrational behavior. The conditions of justifiability and truth of the premises need to be specified prior to this formula being able to do any real work (we may assume that the logical and inferential criteria are the least problematic here). In the realm of theoretical reasoning, the epistemological difficulties of fully spelling out the justification conditions for a premise are well-known. In practical reasoning, the case is further complicated by the unique nature of the major (motivational) premise – it (typically) expresses a desire rather than presents a description of the fact. Whereas in theoretical argument the problem is in determining whether the factual statement is true or false (or at least acceptable, given the context and the justificatory reasons), there is no corresponding problem in case of a practical argument. The main question here (normally) is not whether it is *true* that the agent has this desire or goal, but rather whether he *ought to* have this particular desire (or whether it is rational for him to take on this particular goal).

At this point Audi returns to Aristotle, Hume and Kant and their three respective conceptions of rational action. All three of them recognize that an action can be rational or non-rational precisely in virtue of its link to the purpose or the end of the agent. Rationality of action is (in part) a function of its being conducive to achievement of some end. But whereas Aristotle and Kant put strict constraints on the kind of ends that can normatively ground an action, Hume offers a purely instrumental account of rational action. For Aristotle, happiness is not only the motivational foundation of every rational action (i.e., every action can be traced through a 'purposive chain' to this ultimate end), but it is also the normative ground – happiness is that which one *ought to* pursue, not simply what one *does* pursue as a matter of fact. Given this normative goal, the evaluation of practical reasoning and action

based on it straightforwardly follows: an act X is rational insofar as it "subserves the end of happiness in a way that accords with right reason."[360]

Like Aristotle, Kant postulates the constraints on the acceptable end of action which will determine its rationality. Rational norms of moral reasons take precedence over other reasons in the normative order: it is what the agent *ought to* desire. But he also concedes the motivational power to the requirements of practical reason. As autonomous rational agents, we are not only able to see what we (morally) ought to do, but we can also act on this understanding through our free will.

By contrast, Hume paints a very different picture. Given his famous contention that reason is motivationally inert, it is still open for Hume to label an action rational or irrational. Yet his basis of such evaluation is limited to the instrumental features of action. The means-end rationality of Hume's position takes no notice of the nature of the end of action, but is merely a function of its success to achieve that end. The role of practical reason is limited to determining the optimal way of achieving the desire, but it places no *a priori* constraints on the content of this desire. As Audi remarks, "Hume does not hold that it is rational to desire even pleasure or the absence of pain."[361] Audi defines Hume's position as that of "unrestricted pluralism" – it provides no tools for limiting the set of all rational ends that an agent may assume, and does not discriminate against even the most peculiar desires. Unlike Aristotle, Hume cannot say that an action leading to self-destruction is necessarily irrational (because, say, it is not conducive to happiness), and unlike Kant, he would not (necessarily) classify an immoral action as being contrary to reason, since in both cases a substantive account of the rational object of desire is presupposed.

---

[360] *Ibid*., p. 211.

[361] *Ibid*.

Audi rejects Hume's account, and, in the course of his general criticism of the 'maximization of expected utility' view of rational action (which he sees as being similar to Hume's view), he argues for a position closer to Aristotle and Kant. The Humean account denies that there are any substantive criteria for the rationality of desires (the ends), which means that we cannot assess the *motivational* (major) premise in a standard practical syllogism - and this is clearly unacceptable for Audi. Moreover, he wants to argue that it is sometimes rational for an agent to do that which fails to maximize one's expected utility, i.e., that which is not in one's immediate interest. This idea is familiar enough, and it is based largely on Kantian grounds. The Kantian claim that certain objects of desire are intrinsically valuable and, as such, constitute ends in themselves, is tentatively accepted by Audi but with important qualifications. The most significant modification concerns the extent of those intrinsically valuable objects of desire – according to Audi, we should not limit them to moral values. There are also a number of other desires, such as a desire for the happiness of others for its own sake, which can legitimately be at the end of 'purposive chain' (as an end in itself), and thus can normatively ground rational action. According to this pluralistic conception of intrinsic desires and unconditional values, "it is also rational to want, for its own sake, the flourishing of philosophy, science, and all the arts."[362] Audi's position thus can be best described as a 'modest pluralism' as compared to Hume's 'unrestricted pluralism'.

In effect, Audi argues that the notion of *rationality* and the notion of *well-groundedness* are equivalent. An action is rational if and only if it is grounded in the right kind of way. In this he follows a long tradition in philosophy from Aristotle to Kant which takes the notion of a *good ground* for an intentional action to be a matter of the acceptable

[362] *Ibid*., p. 215.

rational end (motivational premise), undefeated beliefs about the means for this end (factual premise), and its being performed *on the basis* of the reasoning contained in these two premises.

In what follows, I presuppose Audi's account of rationality of action as the most reasonable view on the matter. Its main virtue is the recognition of *plurality* of values (both moral and non-moral) that could normatively ground an action. It further follows that an action based on specifically *moral* reasons is not necessarily the *only* kind of rational action in the circumstances. In certain situations, it might *also* be rational (and thus equally acceptable) to act on non-moral reasons instead,[363] even though acting morally *cannot* be construed as an irrational action.[364]

In order for us to be able to attribute irrationality to amoralists, then, it remains to show that the reasons and values that typically underlie the amoralist's behavior (expressed in the major premise of a practical argument) *cannot* plausibly be construed as belonging the set of all rationally acceptable final goals and ultimate values. Hence the pluralistic conception of values accepted in this work *must* be narrow enough to exclude certain desires and certain ends as irrational and thus unacceptable. The force of this emphatic *must* will be explained in the following section.

3.2 The Case of Ted Bundy

[363] E.g., "If my life would be significantly altered by keeping the promise and you would be only inconvenienced by my breaking it, then my breaking it, while prima facie wrong, may be excusable." (Audi, p. 185). Some of the unpleasant implications of the stronger (Kantian) position which takes moral reasons and moral values as *always* overriding all other reasons and values are explored by Susan Wolf in her article "Moral Saints" (1982).

[364] The point that it is never *irrational* to act morally is specifically argued for by Bernard Gert, "Morality" (1998).

The reasoning of Ted Bundy, one of the most notorious serial murderers in American history, may perhaps illustrate the need for a more restrictive conception of the rationally acceptable values better than strict formal arguments (if there were such) to the same effect. Bundy was an educated (B.A. in Psychology) and charming young sociopath who killed (by bludgeoning or strangulation) and raped at least 30 women and girls, the youngest victim being just 12 years old.[365] In a tape-recorded conversation between Ted Bundy and one of his female victims, Bundy presents a rather sophisticated justification of his behavior, which deserves quoting in full:

> Then I learned that all moral judgments are 'value judgments,' that all value judgments are subjective, and that none can be proved to be either 'right' or 'wrong.' […] Believe it or not but I figured out for myself that if the rationality of one value judgment was zero, multiplying it by millions would not make it one whit more rational. Nor is there any 'reason' to obey the law for anyone, like myself, who has the boldness and daring – the strength of character – to throw off its shackles. […] I discovered that to become truly free, truly unfettered, I had to become truly uninhibited. And I quickly discovered that the greatest obstacle to my freedom, the greatest block and limitation to it, consists in the insupportable 'value judgment' that I was bound to respect the rights of others. I asked myself, who were these 'others'? Why is it more wrong to kill a human animal than any other animal, a pig or a sheep or a steer? Is your life more to you than a hog's life to a hog? Why should I be willing to sacrifice my pleasure more for the one than for the other? Surely, you would not, in this age of scientific enlightenment, declare that God or nature has marked some pleasures as 'moral' or 'good' and others as 'immoral' or 'bad'? In any case, let me assure you, my dear lady, that there is absolutely no comparison between the pleasure I might take in eating ham and the pleasure I anticipate in raping and murdering you. That is the honest conclusion to which my education has led me – after the most conscientious examination of my spontaneous and uninhibited self.[366]

---

[365] Michaud, Stephen, & Aynesworth, Hugh, 2005.
[366] This is a statement of Ted Bundy quoted by Jaffa, Harry, 1990, pp. 3-4.

We may disregard for now the question to what extent the above reasoning process was the *actual* cause of Bundy's life-style as a serial murderer, or whether it was a subsequent *ad hoc* rationalization of his already immoral attitude and behavior. But it seems at least possible that the kind of moral nihilism advocated by Bundy in this passage *was* among the motivating causes of his actions.

There are several points in the passage quoted that merit a detailed second look. First, Bundy presents his thoughts on the rational status of moral judgment as something he *learned* at one point. Even though the exact source of this knowledge is not mentioned, it was this initial intellectual discovery that led to all the subsequent criminal actions. More specifically, he claims to have learned that "rationality of one value judgment was zero," that it was "insupportable," and that there was no reason "to obey the law for anyone, like myself, who has the boldness and daring – the strength of character – to throw off its shackles." We can best understand this claim as the denial that one's final goals and ultimate values are subject to rational appraisal. It is a familiar Humean point which recognizes only instrumental or means-end rationality, but denies that one's desires as such can be rational or irrational. What Bundy discovered, then, was that (roughly) Hume's account of practical reasoning was right.

The rest follows with admirable logical consistency. If a moral normative judgment posits a certain value (e.g., respect for rights of others), which itself cannot be rationally grounded, then the given valuation is in the final analysis *arbitrary* ("insupportable"), and one has no (external) reason to honor this value unless he chooses it as one's own end. Next, as a middle step, if one morally *ought* to do X, then one has a reason for X-ing.[367] Since

[367] This is what Richard Joyce calls "Mackie's Platitude." (2002, p. 46).

Bundy has no reason to respect the rights of others (it is not one of *his* values), it follows that the prescriptive judgment to the effect that he ought to do so would simply not be applicable to his case.

But what *are* Bundy's own values? He mentions at least two in this passage: 'true freedom' and 'pleasure.'[368] Once accepted as the normative ground of his actions, these values determine the rationality of each particular choice. Roughly, a choice is rational to the extent it promotes one of these final ends. He further argues that becoming "truly uninhibited" and "throwing off the shackles of the insupportable [moral] value judgment" are the most efficient means for becoming 'truly free' and experiencing as much pleasure as possible. The constraints of morality, thus, are in conflict with his goals; indeed, the impression is conveyed that the requirements of the moral value-system and those of the pleasure-based system are contradictory.

We should pause here for a second. Assuming, for the argument's sake, the restriction of rationality to means-end domain, one might still question whether Bundy's chosen means (namely, immoral actions) are indeed conducive to his end (say, pleasure). Might he not be able to achieve a greater degree of satisfaction (in the long run) by pursuing a more socially acceptable life-style? If the answer to the last question is 'yes', then we can properly charge Bundy with instrumental irrationality, or, at the very least, with being less then fully reasonable in his choices and actions.

My contention is that there is no reason for supposing Bundy being less then fully competent at instrumental reasoning that determined his specific choices. The above objection presupposes that 'pleasure' is a univocal term that denotes the experience

[368] It might be argued that Bundy's *final* end is pleasure, and 'freedom' (in his sense) has only instrumental value, in so far as it is conducive to pleasure. But nothing important hinges on this point. In any case, there is no incongruity in postulating two or more final ends.

essentially the same in kind in all pleasure-giving activities. However, when Bundy claims that "there is absolutely no comparison" between the pleasure of eating ham and the pleasure of raping and murdering a woman, he is, in effect, endorsing something close to an Aristotelian account of pleasures, according to which pleasures differ 'in species' depending on activities that they crown.[369] It would be odd to suppose that the intensity of pleasure Bundy feels at, say, murdering a woman, might be eventually 'outweighed' by increasing the amount of ham that he can eat for free. Hence, it is not pleasure *sans phrase* that he pursued as his goal, but pleasure of a *certain kind* – the one, which, presumably, could only be experienced by doing criminal actions.[370] Granted that a person himself is the best judge of which activities are most pleasant (given his tastes)[371], it is desperate to attribute to him some obvious mistakes in instrumental reasoning. Indeed, his actions are in many respects the paradigm examples of rationality and efficiency. His logical skills are impeccable, his abilities to reason abstractly and learn from experience are admirable, and the level of his general reflective intelligence seems to be higher than average. Ted Bundy is an example of a perfectly *rational* amoralist.[372]

To return to our analysis of Bundy's statement, we should finally look at this provocative question of his: "Why is it more wrong to kill a human animal than any other

---

[369] *Nicomachean Ethics*, 1175a25ff.

[370] For brevity sake, I shall omit the discussion of whether *freedom* (Bundy's other value) could be achieved through other, non-criminal life-style. It appears that his conception of freedom is close to Hobbes' notion of the natural liberty (*jus naturale*) that humans have in the state of nature. If that is so, then, indeed, preserving this unrestricted freedom is incompatible with any normative constraints.

[371] Aristotle's 'objectivist' account of pleasures notwithstanding (*N. Ethics*, 1176aff). My main reason for rejecting Aristotle's contention that pleasures of a vicious man are not *really* pleasures, and that it is only the virtuous person's experiences that should be taken as the criterion in identifying something as pleasant have to do with the obvious problems of applying the same 'objectivist' reasoning to the experience opposite of pleasure, namely *pain*.

[372] To the extent, of course, to which *perfection* in rationality is humanly possible.

animal, a pig or a sheep or a steer?" If wrongness is defined against the background of means-end normativity, then it follows that a 'wrong' act is the one that fails to lead to the chosen goal, just like a wrong road is the one leads away from the desired destination. Since Bundy's 'destination', as we have seen, is a certain kind of pleasurable experience,[373] which (presumably) can only be achieved by doing criminal action, including murder and rape, then there is literally *nothing wrong* for Bundy to kill 'a human animal.'

I take the reasoning in the last paragraph as one of the clearest cases of *reductio ad absurdum*. If, given a certain conception of rational action, the actions of Ted Bundy and other amoralists turn out to be fully justified from the point of view of rationality, then there *must* be something wrong with the conception itself. And in fact, there is. The mistake lies in the assumption that the motivational premise of the practical argument (e.g., "I desire to **φ**.") is beyond appraisal in respect of its rational status. Setting specific constraints on what kind of things can be *rationally* desired (as one's final goals) is the only way to avoid the conclusion of the last paragraph. As I have suggested in the previous section, Robert Audi's theory of 'modest pluralism'[374] can and should be construed in a way that excludes certain goals (e.g., unrestricted pursuance of pleasure) from the plurality of all rationally acceptable ends, even if it recognizes other ultimate values in addition to the moral ones.

Providing the complete list of all rationally acceptable values, together with the proper hierarchy within the class, is a task for a different project. It is enough for the present purposes to postulate that, from the rational point of view, the final ends of action *cannot* be such that the action-guiding principles properly adopted for their achievement are strictly

---

[373] Cf. Wolman's claim that "the 'pleasure principle,' that is, the principle of immediate gratification of needs, is the main motive in the life of sociopaths." (1987, p.44)

[374] By contrast with the Humean 'unrestricted pluralism.'

incompatible with the principles sanctioned by the ends of morality. The incompatibility that is ruled out by this claim is 'strict' in the sense that it refers to the relation of logical contradictoriness between the two sets of principles, but it allows for milder forms of logical conflict.[375] Hence, we can repeat what was said above: it is never *irrational* to act morally even if it is sometimes rational to prefer a non-moral value over the moral one. Still, the intuitive superiority (although *not* indefeasibility) of the moral values over the non-moral ones should ensure that the circumstances where an action contrary to the requirements of morality is (rationally) called for are relatively rare. As Audi puts it, "what we may conclude about rational conduct in relation to moral conduct is that a rational person, when suitably informed and adequately experienced, will *tend* to act morally."[376] By contraposition, a person's tendency to act immorally (where this is more than atypical moral lapses) does bespeak deficient rationality.

## 3. Summary and Conclusions

This chapter has explored the various senses in which an amoralist might be said to fall short of the demands of cognitive excellence.

I have argued that, given the theoretical reasons and the factual evidence cited above, it is difficult to maintain that the persistent moral failures of amoral individuals are due to intellectual incompetence of some sort, including the inability to master the moral concepts or else, understand certain morally relevant facts (e.g., the fact of the reality of other people).

---

[375] A value system *A* which sanctions a principle "S ought to do X" is strictly incompatible with a value system *B* which sanctions a principle "S ought to do non-X". The milder forms of logical conflict would be exhibited by the two value systems issuing the following two principles correspondingly: "S ought to do X in circumstances C" and "S ought to do Y in circumstances C."

[376] Audi, 2006, p. 185. [Audi's emphasis].

Nor is it plausible to view amoralists as suffering from moral ignorance or any exceptional failures of instrumental, means-end rationality.

These results forced us to take a closer look at the suggestion that the kind of cognitive failure involved in most cases of amoralism is a mistaken value judgment. Audi's account of practical argument and his defense of the 'correspondence thesis'[377], firmly rooted in tradition, provided a theoretical ground for the possibility of evaluating one's intentional action from the rational point of view, even in cases where no explicit practical reasoning was taking place. It was further argued that both the major premise (the motivational premise expressing a desire or postulating a value), and the minor premise (a belief specifying the means for achieving the end) are subject to rational evaluation. The 'modest pluralism' of values, which recognizes both moral and non-moral (final) ends of action, was adopted as the most plausible view on the matter. Nonetheless, however pluralistic, the conception must be demanding enough to exclude from the set of all rationally acceptable ends, the values and desires characteristic of the amoral individuals, like Ted Bundy.

The moralist of the Kantian mold may see it as an obvious fact that a rational human nature has unconditional worth which demands respect and restricts the range of possible behaviors toward other humans, whereas Ted Bundy prizes his own pleasure and freedom from constraints above all and apparently finds nothing especially valuable in human rationality or autonomy. The two sides differ in their conceptions of what is valuable, of what is good and evil, and from the perspective of the account defended here this difference may best be explained by postulating that the conclusions of Ted Bundy, however 'honest' and 'conscientious', rest on a mistaken (indeed, irrational) premise.

---

[377] I.e., the view that to every intentional action there corresponds at least one practical argument whose premises express motivation and belief jointly sufficient to explain the action. (Audi, 2006, p. 103)

Bundy's hedonistic ethical egoism (if we may put a label on his worldview)[378] led to a lifestyle that may be properly described as amoral (as well as immoral). In his case, however, there is no reason to suppose that a person like Ted Bundy has a diminished capacity for role-taking or fails to understand the meaning of certain moral terms. Rather, his system of values and the adopted hierarchy of (non-moral) normative principles simply fail to give any significant weight to the feelings and interests of others even if he, just like the sadist mentioned above, might still be able to recognize those feelings and interests as quite real. If my argument in this chapter is sound, we should also be able to say that he *should have* assigned significance to the interests of other human beings, and that his failure to do so is indicative of a *mistaken* or *irrational* judgment about the hierarchy of values.[379] Thus, a mistaken metaethical judgment about the proper place or objective status of certain values appears to be the best candidate for the underlying cause of amoralism.

In this chapter I have tried to provide an answer to the question "*What* is wrong with amoralists?" I have argued that a failure to recognize certain morally relevant facts or values as sufficient reasons for actions (including the actions of omission) is what ultimately leads to amoral condition. This conclusion partly satisfies the quest for explaining the amoralist. Yet, an explanatory project should not stop with the answer to the 'what'-question, but must, when possible, pursue the further 'why'-question as well. Hence, in the final chapter, I shall

---

[378] We may recall here (from Chapter Three) Richard Jenkins' claim that "overwhelming egocentricity" is the key character trait of all psychopaths (1960, p. 324).

[379] A plausible account of the *sense* of this categorical 'should' is given by Michael Smith in "The Moral Problem" (1994). Smith's position has much to recommend. According to his view, the best way to construe the objective status of normative reasons is in terms of the agent's *own* potential desires and motivating states when the agent is considered in an epistemologically privileged situation (e.g., as fully rational with all and only relevant beliefs). Unlike the purely Kantian view, or the 'Ideal Observer' account of moral reasons (e.g., by R. Firth), Smith's position tries to avoid the danger of *alienating* the agent from the normative reasons that make demand on his actions. After all, it is what the 'improved version' of the agent *himself* that would have those desires.

focus on the question "*Why* do amoralists fail to recognize certain values?" and explore the limits of our ability to arrive at the correct conclusion.

# CHAPTER VI

## RESPONSIBILITY AND THE LIMITS OF EXPLANATION

> Evil ever remains man's own choice; the ground cannot cause evil as such, and every creature falls through its own guilt. But just how the decision for good or evil comes to pass in the individual man, that is still wrapped in total darkness.
>
> F. W. Schelling, *On the Essence of Human Freedom.*

In the preceding chapters I have argued that in approaching the problem of amoralism it is much more promising to focus on the cognitive failure of these individuals rather than on their emotive deficiency. It was suggested that the kind of cognitive mistake involved in cases of amoralism is a mistaken judgment about the status and the proper place of moral values. A mistaken conception of the appropriate (final) ends of one's choices characterizes both the extreme and selective amoralists. We have seen that the mistake that the amoralists make need not be perceived *as* a mistake by the amoralists themselves. There is no need to suppose, furthermore, that this realization *would* come to an amoralist in some more favorable (but practically feasible) epistemic state. There is a legitimate sense in which one

can be under normative obligation to respect a certain end (i.e., have a *normative* reason for action), without being open to the possibility of ever recognizing this obligation or being moved by it.[380] In so far as we are staying within the sphere of experience, we may acknowledge that in many cases no amount of rational persuasion would change the amoralists' outlook, and even the strongest admonitions would fall on deaf ears. One's conception of ultimate values constitutes one of the most fundamental levels of practical reasoning, and, once solidified, is quite resilient to any radical changes.[381]

In what follows I shall directly address the question *why* some people refuse to acknowledge the authority of moral constraints, but choose a non-moral action-guide instead. I shall argue that the prospect of finding the answer to this 'why' question stands in direct relation to our willingness to hold amoralists morally responsible for their actions, and the plausibility of a particular account of free actions.

## 1. The Question of Responsibility

If my argument in the previous chapter is sound, the cause of the mistaken value judgment can be attributed neither to the deficiency of general reflexive intelligence of amoralists nor to (excusable) ignorance of the content of the moral rules and values. There are compelling reasons to believe that amoralists are not disadvantaged in these respects as compared to normal subjects. I shall argue below that the mistake in question is the result of a conscious choice on the part of amoralists, in a situation where different choices were

[380] This *externalism about reasons* is convincingly defended, for instance, by Russ Shafer-Landau (2003; pp. 176-88). Cf. also B. Williams' claim: "The moral law is more exigent than the law of an actual liberal republic, because it allows no emigration" (1985, p. 178).

[381] It is widely accepted that all of the existing clinical treatments aiming at reforming the attitudes of psychopaths or correcting their behavior over the long period of time are largely unsuccessful (see Cleckley, 1982; Wolman, 1987). Cf. also Aristotle: "The person who is now unjust or intemperate was originally free not to acquire this character, so that he has it willingly, though once he has acquired the character, he is no longer free not to have it now." (*Nicomachean Ethics*, 1113b 20-23)

always possible. The decision to pursue a particular end incompatible with the ends of morality need not be consciously made at some point in time. The actions (but not necessarily the words) of amoralists can and must be interpreted as the best indications of which values *actually* guide their choices.

It appears, in my judgment, that we can best approach the explanatory project under consideration from the side of the question of moral responsibility of amoralists. More specifically, I shall defend here the soundness of the following deductive argument:

1. If amoralists are morally responsible for their choices and actions, then their behavior cannot be fully explained.
2. Amoralists are morally responsible for their actions.
3. Therefore, their behavior cannot be fully explained.

The argument is an instance of a valid *Modus Ponens* form, which implies that the truth of the two premises would logically necessitate the truth of the conclusion. It remains then to examine whether the two statements that appear in the premises are indeed true. I shall examine each of these statements separately, starting with the second one.

1.1 How Different the Amoralists Are?

My main reason for holding extreme amoralists morally responsible breaks down into two claims: the affirmation that most ordinary people *are* morally responsible for their intentional actions and the denial that the psychopaths (and other amoralists) are *qualitatively* different from the non-psychopaths. Even though emphasizing the distinction between amoralists on the one side and moralists on the other did help to bring to the fore the characteristic features of these amoral individuals, it is highly questionable whether the difference between the two classes is fundamental enough to warrant a differential

application of the demands of responsibility. It was stressed from the beginning that amoral condition is a matter of degree, and it can vary significantly both with regard to its *scope,* and its *intensity*. We have seen that individual cases may exhibit various degrees of moral unresponsiveness, ranging from partial and temporal exclusion of a certain group of people from the domain of all moral subjects (as regularly occurring during military conflicts), to more persistent cases of selective amoralism (e.g., Eichmann), and culminating in extreme instances of thoroughgoing moral apathy (psychopaths; serial killers). Likewise, one can be more or less indifferent to the universal ends of morality, and the strength of one's moral motive need not remain the same over time. Thus it appears that there exists a *continuity* between the 'normal' subjects and amoralists, including the most radical cases of psychopathy. If it is true, then, that there are no good reasons for viewing psychopaths as a different *type* of agents, then there are no reasons for treating them differently with respect to moral accountability.

The above thesis, however, needs to be further substantiated. A number of researchers who want to exempt extreme amoralists from responsibility (both moral and legal) build their case on the assumption that psychopaths are *too* different from the rest of us to warrant the application of the familiar moral and legal categories. We have seen in Chapter One that Aristotle introduced a category of a 'brute' to refer to the most perverted kind of agent whose wickedness goes "beyond the human level".[382] Given his graphic description, the impression is conveyed that a 'brutish type' is indeed a qualitatively different type of agent – a kind of a borderline case between humans and animals. A contemporary philosopher Jeffrie Murphy is even more overt when discussing the status of psychopaths: "From the moral point of view, it

[382] *Nicomachean Ethics*, 1149a16-17.

is very implausible to regard [psychopaths] as *persons* at all… A psychopath is more profitably pictured as an *animal*."[383] Similarly, Harry Frankfurt (1971) would deny the personhood to psychopaths on the ground that they (apparently) lack the second order desires and thus belong to the class of *wantons* instead.[384] It is these radical conceptions that I want to question.

Let us begin with the simpler case first, namely, with a selective amoralist. One of the persistent themes in Arendt's account of Eichmann's trial is the apparent incommensurability between Eichmann's rather ordinary, shallow personality and the horrendous consequences of his actions. As Vetlesen observes, "Eichmann left Arendt with a puzzle: the actual monstrous consequences of the measures to which the SS officer Eichmann had contributed seemed to stand in flagrant contradiction to the apparent harmlessness of his personality."[385] Eichmann was clearly not 'a devil incarnate'; he cannot be plausibly seen as an inherently evil or beastly person, despite the horrific aftermath of his activity. The gap between the doer and the deed is almost unbridgeable in his case. By all accounts, he was a regular person who abdicated his will to make moral choices, and systematically and routinely gave preference to the values of a "bureaucrat" in the SS system instead. According to Arendt's account, Eichmann epitomizes the 'banality of evil'; his case forces upon us the uncanny realization that mere thoughtlessness and unconditional compliance together with such trivial motives as a desire for promotion or a desire to do one's job well, are often sufficient to bring the enormous amount of evil and suffering into the world.

---

[383] Murphy, 1972, pp. 294-3.

[384] It must be noted that despite this restriction of the category of *persons* Frankfurt still believes that wantons can be held responsible for their behavior, even if they could *not* have done otherwise.

[385] Vetlesen, 1994, p. 86.

There is no need to go into greater details or cite more examples here. It appears obvious that, *if* anyone is responsible,[386] a selective amoralist is just as responsible for his actions as any ordinary person who commits a crime only occasionally or once in a life-time. If a person responsible for the wrongful death of one individual is guilty of a crime and could be held morally and legally accountable, then also a person who contributed to the death of millions is equally fit to be treated as a responsible moral agent, granted there are no relevant psychological differences between the two cases (e.g., one is not a psychotic). The last condition has been shown to obtain in the case of Adolf Eichmann, but can perhaps be equally well satisfied in most other cases of selective amoralism. Hence, there is no good reason to suppose selective amoralists to be incapable of standing the trial of the moral or legal law.

Admittedly, the case of the extreme amoralists is much more controversial, and the disagreement about the extent of responsibility of various sociopaths has been deep ever since this topic began to interest researchers and the law enforcement officials. A number of authors have denied responsibility to psychopaths on the basis of their marked difference from the rest of us, a difference both in values and patterns of thinking. Thus Anthony Duff expresses a common enough attitude when he writes:

> A psychopath is seriously defective in practical understanding and rationality. […] This deficiency is clearly a *disorder*: a psychopath cannot understand the nature and quality of his actions: he cannot control his actions in the light of any rational concerns or values, not because his impulses are strictly irresistible, but because he has no conception of rational values as providing reasons for actions. We cannot hold him answerable for his actions, any more than we can a young

[386] I am assuming here that a certain interpretation of a hard deterministic position which denies moral responsibility altogether and universally, is false.

child.[387]

These comparisons of a psychopath with a young child or even with an animal (as in Murphy) are meant to suggest a type-difference between the psychopaths on the one hand and the non-psychopathic individuals on the other. This difference, in turn, should warrant the difference in treatment in a number of social aspects, including the exemption from responsibility.

The defenders of the above position rest their case on a questionable assumption that there exists a sharp and stable divide between psychopathic and non-psychopathic personalities. Yet, the available clinical and physiological evidence clearly favors the view of the *continuity* between the two classes, rather than the type-difference. Thus the philosopher/psychiatrist Karl Jaspers sees psychopathic behavior as only exaggerated extension of the normal personality, and not abruptly discontinuous with it.[388] He observes that with the cases of anti-social personality disorder, unlike the cases of the major psychoses such as schizophrenia, we are mostly limited to observing the behavioral manifestation of this condition, and there are little, if any, reliable internal physiological ('somatic') or invariable psychic symptoms. In a similar vein, psychiatrist Richard Jenkins argues (contrary to Duff) that the psychopathic personality "is *not* so much a disorder as it is a defect in development. In one sense we are all born psychopaths. Most of us outgrow this stage of overwhelming egocentricity: some become arrested there."[389] Egocentrism to a greater or lesser degree is certainly not a rare trait in many people, and it would be odd to suppose that

---

[387] Duff, A., 1977, p. 199; The incomplete list of philosophers and psychologists who argue for a similar thesis, denying the responsibility of psychopaths, would include Haksar, (1964 & 1965); Murphy, (1972); Pritchard, (1974); Strawson, (1974), Arrington, (1979), and Deigh, (1995). For the opposite view see Frankfurt (1971); Flew, (1973); Smith, R. (1984); and Glannon, (1997).

[388] Jaspers, 1963, pp. 724ff.

[389] Jenkins, 1960, p. 324.

by increasing the scope and intensity of this attitude one would eventually pass over to a qualitatively different type of agent, where none of the normal standards of responsibility are applicable. As Jenkins reminds us, "the psychopath is simply the extreme on the continuum. This means that the significant question will be one of degree rather that one of categorization, not, *is* he a psychopath, but rather *how* psychopathic is he?"[390]

Robert Smith cites further evidence showing that the best attempts to correlate psychopathy with abnormal electroencephalogram (EEG) readings (the view that was enthusiastically pursued in the early stages of research) have failed. Despite the promising beginnings, no unique or identifying brain wave pattern has emerged among diagnosed psychopaths. Gale, for instance, estimates that as high as 25 percent of unselected youth samples show atypical EEG patterns, and that such patterns are common among many psychiatric disorders.[391] Likewise, deFundia, Draguns and Phillips argue that "pathological behavior in general represents merely the caricatured extreme of culturally shared modes of behavior, and this would be particularly the case for psychopathy because of the social nature of the symptoms."[392]

Finally, a growing number of criminologists are wary of the once standard distinction between an ordinary criminal and a psychopath. The opinion of Stanton Samenow is worth citing again in this context: "Although diagnosticians may make distinctions between the psychopath and criminal, for all ostensible purposes, one differs hardly at all from the

---

[390] *Ibid*. Cf. also Jeffrie Murphy's claim that "we all have our psychopathic tendencies, and so all the actual cases may be neither black nor white but various shades of gray." (Murphy, 1972, p. 296).

[391] Smith, R., 1984, p. 184.

[392] *Ibid*., p. 184-5.

other."[393] On this view, being more persistent in crime and less susceptible to correction does not yet put a person in a radically different category from the rest of criminals.

In the absence of sufficient evidence to the contrary, we are entitled to conclude that the amoralists are essentially similar to the normal subjects in all aspects relevant for being morally responsible even if they have radically different values or priority of values. That is, when authors like Pritchard describe psychopaths "as though they are from another world, another 'form of life'," [394] they let themselves be overly impressed by the external differences in psychopaths' values, beliefs and motivating goals, but tend to overlook the same basic underlying structure of purposeful behavior that they share with others, and which alone is relevant for moral responsibility. We have also seen in Chapter Four that the well-documented difference in emotive responsiveness of psychopaths as compared with normal subjects (the difference that, we may agree, is genetically predetermined) does not yet condemn the 'insensitive' person to amoral life-style (as autistic individuals show). At the end, it is one's own free choice rather than one's 'nature' that commits him to a life devoted to values inconsistent with values of morality. [395]

None of the above is meant as a denial of usefulness of separating psychopaths into a separate category – it is often helpful to isolate one segment of the continuous scale in order to bring the phenomenon into a greater focus, but it is important to remember that the classification does not always correspond to a different 'natural kind' with unique properties. An unrelenting racist, rational egoist and a violent serial murderer are all possible positions

[393] Samenow, 1984, p. 181.
[394] Pritchard, 1974, p. 642.

[395] The exception should be made for cases where the amoral state was caused by brain damage (e.g., the case of EVR cited in the previous chapter). In those cases, where there *is* an identifiable physiological cause of the anti-social behavior, the agent should be exempt from responsibility.

on the same 'line' with greater or lesser proximity to the extreme points. This etiological continuity between the normal and amoral subjects allows us to extend the usual demands of moral responsibility to all entries on this imaginary scale.

### 1.2 Glannon and Partial Responsibility

It remains now to examine a view that seeks a middle ground between the two extreme positions. Walter Glannon in his relatively recent contribution to the debate makes a strong case against Murphy, Deigh and other philosophers who want to exempt psychopaths from responsibility altogether, but stops half-way when he argues that psychopaths are only *partly* responsible for their criminal actions. This view deserves a closer analysis.

Regarding the necessary conditions for being morally responsible, Glannon accepts the following thesis: "A person is an appropriate candidate for attributions of moral responsibility if and only if he is capable of making a certain response to moral reasons."[396] (The list of '*moral* reasons' would minimally include the needs and interests of other people.) This general condition is in turn analyzed into two distinct capacities: (1) The capacity to recognize that others have needs, interests, and rights; and (2) the capacity to be motivated to act on these reasons. In the rest of the article Glannon seeks to establish that psychopaths do have the first capacity, but lack the second capacity – they are inherently unable to respond to moral reasons properly. Thus, he concludes, since only one capacity out of two is present, psychopaths are only *partly* responsible for what they do.

My contention is that Glannon's argument, although original and stimulating, does not withstand closer scrutiny. We may agree with Glannon that both capacities mentioned are

[396] Glannon, 1997, p. 263.

needed in order for an agent to be an appropriate candidate for moral praise and blame. Furthermore, as I have argued in previous chapters, psychopaths *are* indeed capable of recognizing the presence of the morally relevant facts (e.g., pain, desires, needs). They do not generally suffer from a cognitive impairment that would prevent such recognition. However, it is the denial to psychopaths the capacity to be appropriately *motivated* by those reasons that I want to question.

To begin with, how could this "incapacity to respond to moral reasons" be established? Presumably, using the observable behavioral criteria only. To clarify this, we may imagine two individuals, Steve and John. Steve was never observed to respond to moral reasons appropriately; e.g., he passed by five different swimming pools with drowning children, and never bothered to do anything to save the children. John, on the other hand, was only rarely moved to save children in similar circumstances; we can assume that, typically, he would save only one child out of five in the circumstances cited. Given this behavioral data, should we count John, the person who occasionally saves a drowning child, as incapable of responding to moral reasons? Surely not, since he is motivated appropriately at least sometimes. Should we say that Steve, on the other hand, lacks the capacity to respond to reasons of morality? Not necessarily. We cannot infer from the fact that Steve was never *observed* to act morally, to the stronger claim that he is *incapable* of doing so.

Even if we grant, for the sake of the argument, that a lack of behavioral manifestation is indeed a reliable indication of the inner incapacity, still, we may question whether a person like Steve ever existed except in a philosopher's imagination. How plausible is it to suppose that a psychopath *never* acted morally (i.e. for moral reasons), that he never kept a promise or

helped a stranger, even as a child?[397] It might be true that a genuine psychopath never gives priority to moral reasons in cases of conflict with his own values, but, in the absence of the opposite incentives, there is nothing that would prevent him from responding appropriately to the demands of morality. After all, he is not a person who made it his chief goal in life to act contrary to the demands of morality at all costs – rather, he is largely indifferent to these demands. But even a single instance of moral reaction would be sufficient to establish the *capacity* in question.

Furthermore, Glannon's theory commits him to a counterintuitive view that John, who only occasionally acts morally, is *more* responsible and hence *more* guilty than Steve, who never does. In fact, it appears that all that a criminal needs to do to avoid being held responsible for his immoral actions is to act immorally from now on without a single exception. As Gary Watson rightly observes, the view that extreme evil disqualifies one for blame is "paradoxical."[398]

Lastly, and more seriously, Glannon's theory of partial moral responsibility of amoralists rests on the assumption that responsibility admits of degree, i.e., that one can be more or less responsible for his voluntary actions. This is by no means obvious. We should recall that Glannon builds his case for partial responsibility on the crucial claim that the two necessary conditions for being responsible are only *partially* satisfied in psychopaths. But this reasoning relies on the general assumption that a partial satisfaction of the necessary conditions for a given function result in *partial* functionality. Let us suppose that in order for a car to be able to move the two conditions need to be met: a working engine and there being

---

[397] To argue that one is not acting *qua* psychopath in those rare moments when he *does* respond to moral reasons is effectively to include the relevant incapacity into the *definition* of a psychopath, and thus beg the question against the alternative view.

[398] Watson, 1987, p. 268.

enough gasoline in the gas-tank. The car does not become a *partly* moving car, however, if only one of those conditions obtain. Glannon needs to show, in other words, that the case of moral responsibility is essentially different from the car-scenario above.

But let us grant that the notion of partial responsibility is not incoherent. Still, how should we understand this claim? Arguably, he might have meant one of the two things here: **(a)** psychopaths deserve some lenience and should not be held wholly accountable for their actions (e.g., should not be prosecuted to the full extent of the law), since certain extenuating conditions (e.g., non-blamable ignorance) are present whenever a psychopath acts. Or: **(b)** Due to some peculiarity of the psychopaths' nature, they are inherently *less capable* of being responsible for their actions than other agents, even if they are still capable of being partly responsible.

I shall start with the interpretation **(b)** first. It is not clear whether this reading can be squared with the common intuition. In the legal system, one can be partly *guilty*, in the sense that he is guilty only on some of the charges that were originally presented against him. To be partly guilty amounts to being guilty of doing X, but not being guilty of doing Y. On the other hand, if it was established that one has committed X, then the agent is either responsible for his action or not (e.g., by reason of insanity). To be sure, there might be various extenuating circumstances that would diminish the agent's guilt, but those excusing condition do not take from him the *capacity* of being responsible in the first place. A criminal does not become *less* capable of responsibility simply because, for example, his family circumstances warrant a less severe sentence from a judge. The capacity itself, it appears, is either present or not, and cannot be augmented or diminished depending on the external circumstances.

But perhaps Glannon had the interpretation **(a)** in mind. In that case, the point is general enough and it should cover all agents, psychopaths or not: if a person acts in the presence of an extenuating condition (e.g., ignorance), then he deserves lenience. Then it is an empirical question whether psychopaths, as a class, are always acting in ignorance or in the presence of some other extenuating yet not fully excusing condition. If this reading is closer to Glannon's intent, then he must specify the universally present condition which party exempts a psychopathic agent from responsibility for his actions. And in fact, he does:

> Psychopaths may not be responsible for the *consequences* of their actions if their capacity for planning and decision-making is so impaired as to preclude them from foreseeing the likely long-term consequences of what they do. […] Thus while they are at least partly responsible for their actions, they may not be responsible for the consequences of their actions.[399]

First, we may notice that Glannon seems to have changed his initial view here. In the beginning, he argued that what *makes* the psychopaths less responsible as compared to the ordinary criminals is the fact that they are incapable of being appropriately *motivated* by moral reasons. In the passage just quoted, he suggests that the reason why psychopaths should be exempt from full responsibility is that they are incapable of foreseeing the consequences of their actions (i.e., some kind of ignorance, a cognitive deficiency). And it has not been shown that the two positions are consistent with each other or imply one another.

Secondly, this explanation still leaves a number of practical questions open. We may legitimately ask what does it *mean* that a person is responsible for his *actions* but not for the *consequences* of his actions? What counts as a consequence in a given case? Suppose that Ted Bundy's act of killing a victim may (logically) be analyzed into at least two components:

---

[399] *Ibid*., p. 272.

thrusting a knife into his victim's chest (Bundy's action proper) and the victim's death as a direct *consequence* of this action. Should we say that Ted Bundy is responsible for thrusting a knife but *not* responsible for the victim's death? Surely, this would be a rather idiosyncratic way of describing the situation.

If, on the other hand, Glannon does not mean something as immediate as the death of a victim resulting from a hit with a knife, but the *long-term* consequences of one's actions instead that the psychopaths are allegedly incapable of foreseeing (and thus should not be held accountable for), then this empirical claim is very likely to be false. The capacity to foresee the future outcome of one's behavior (with a greater or lesser precision) is (in part) what it means to have an unimpaired general intelligence, and in the earlier chapters we have seen no reason to deny this to psychopaths. They appear to be just as capable as the normal subjects to learn from experience and to make reliable causal inferences about the future states of affairs based on the known evidence about the present. John Deigh's observation is especially relevant here:

> Though amoral the psychopaths appear nevertheless to be capable of reasoning, weighting evidence, estimating future consequences, understanding the norms of their society, anticipating the blame and condemnation that result from violation of those norms, and using these cognitive skills to make and carry out their plans.[400]

We may conclude that the arguments denying the responsibility of extreme amoralists, or trying to limit it in some way, do not withstand a critical examination. Hence, we have no reason to view psychopaths as deserving a special treatment when it comes to moral or legal responsibility for their actions.

---

[400] Deigh, 1995, p. 743.

## 2. Freedom of Will and the Limits of Explanatory Chain

In what follows I shall argue that the behavior that falls under the demands of moral responsibility cannot be fully explained. This position presupposes a certain view on the connection between moral responsibility and freedom of will, as well as a certain conception of freedom of will itself. More specifically, it presupposes that moral responsibility is *incompatible* with causal determinism of one's behavior, and assumes that the essence of a free act or a free choice lies in the fact that it cannot be fully explained or predicted[401] by reference to the antecedent causal conditions.

### 2.1 Van Inwagen's Argument for Incompatibilism

Assuredly accepting the thesis of moral responsibility of amoralists does not yet logically commit one to the acceptance of the position sketched in the previous paragraph. Antony Flew, for one, maintains that psychopaths are not lacking in responsibility, but he argues for this view from the standpoint of *compatibilism*.[402] According to Flew, a psychopath can be held responsible since he is not acting under inner compulsion or under irresistible desires (as a kleptomaniac allegedly does) or under any external constraints, but rather acts voluntarily and "in character." But the emergence of psychopathic character can in turn be fully accounted for (at least in principle) by reference to genetic predisposition, environmental factors or history of reinforcement. There is an unbroken deterministic chain of events, according to Flew, that leads to the formation of one's character, including the adoption of certain values and forming of certain preferences, but the truth of determinism does not rule out moral responsibility. Hence, the inexcusability of one's criminal behavior

[401] I take the notions of predictability and explanation of behavior to be closely connected (cf. Taylor, 1966). Strict predictability implies the possibility for exhaustive explanation and vice versa.

[402] Flew, 1973.

does not preclude, on this model, an exhaustive explanation of one's specific actions or choices.

This is not the appropriate place for a substantial criticism of the compatibilist model of moral responsibility (which traces back to Hobbes[403]), nor can we spend enough time developing an alternative view of freedom of will. For a criticism of compatibilism the reader will be referred to an excellent discussion in Peter van Inwagen's "Essay on Free Will" (1983) and I shall take the liberty of summarizing and largely appropriating the results of that project without attempting to provide any detailed assessment of its success. On the positive side, the account of free will and responsibility that appears most plausible to me, and the one that will be assumed here, is the agent causation libertarianism rooted in the Aristotelian and Kantian traditions and recently defended, among others, by Roderick Chisholm and Richard Taylor.[404]

Van Inwagen essentially argues for the following three claims: 1. People are morally responsible for their actions at least sometimes. 2. Moral responsibility presupposes that an agent acts freely (i.e., that he has *free will*); 3. Free will is *incompatible* with the deterministic picture of human agency.

The first thesis is accepted largely on intuitive grounds – it is a central part of our conception of ourselves as rational agents that we can and should be held accountable for our behavior, and the outright denial of moral responsibility "would be absurd." The second thesis is argued for at some length against the alternative interpretations of moral responsibility which do not list free will of an agent among its necessary conditions (e.g., as

---

[403] See "Of Liberty and Necessity" (in *British Moralists: 1650-1800*. ed. by Raphael, Hackett, 1991, vol. i.)

[404] Chisholm, (1964), 1997; Taylor, 1966.

defended by Harry Frankfurt (1971)). Ordinary thinking about moral responsibility is on Inwagen's side here. We tend to believe that our action is our responsibility *because* our action is something that we determine for ourselves. After carefully examining and criticizing the arguments to the contrary, Inwagen concludes that the ordinary commonsensical view on the relation between responsibility and free will is yet correct: "To be morally responsible for some act or failure to act is at least to be able to have acted otherwise, whatever else it might involve; to be able to have acted otherwise is to have free will."[405]

Finally, Van Inwagen builds a strong case for the *incompatibility* of free will and the view that all our future actions are causally necessitated by the antecedent states of the physical universe. One of his reasons for embracing incompatibilism is that the premises of the arguments for incompatibility of free will and determinism are intuitively more plausible as compared with the premises that function in the best available arguments for compatibilism.

The above position commits Inwagen to some sort of *indeterminism* – the claim that human choices and actions are not wholly necessitated by the past states of the physical universe.[406] Indeed, he cites evidence from quantum mechanics that (on some interpretations) undermines the strictly deterministic picture of the world. However, Inwagen's position does not rest solely on the truth of the thesis of quantum-mechanical indetermination. Rather, in his final analysis, he is arguing for the following inclusive *disjunction*: "we have properties that do not supervene upon the properties of the atoms that we consist of" *or* "free will

[405] Van Inwagen, 1983, p. 162.

[406] Denying the truth of determinism, according to van Inwagen, is not the same as denying the truth of the thesis of Universal Causation.

involves quantum-mechanical indetermination."[407] Free will and responsibility will be safeguarded, according to Inwagen, if at least one of these disjuncts is true.

I conclude that there are good reasons for agreeing with Inwagen on this issue, and thus rejecting both the premise of strict determinism and the compatibilist account of free will and responsibility. In the following section, I shall sketch a positive account of the nature of free action based on the works of Roderick Chisholm and Richard Taylor, which, in my view, best accords with the available evidence and common sense.

2.2 Agent Causation Libertarianism

If amoralists are responsible in virtue of having free will, and freedom of will is incompatible with strict determinism, then there remain only two possible descriptions of the dynamics of a free action. First, one might identify a free action with an *uncaused* action *simpliciter*. Secondly, one might propose an alternative view of causation involved in free action which avoids the first option but safeguards genuine freedom of the agent. The first option leads to a paradoxical result that an agent has nothing to do with the origin of his actions. An action done freely, on this view, is a *random* event whose occurrence could be neither causally explained nor predicted. More importantly, it is not clear in what sense a person could be held responsible for the occurrence of this uncaused motion, if it is not determined by an agent himself. This interpretation of freedom of will views free actions as something that *happens* to a person rather than something that a person actively does, and, hence, it is incompatible with the view that we are morally responsible for our choices.

The above reasoning forces us to look for an alternative account of free action. Roderick Chisholm draws on a long philosophical tradition which goes back to Aristotle

[407] *Ibid*., p. 217.

when he defends a view that avoids the pitfalls of both strict determinism and the various indeterministic explanations of a responsible action:

> We must not say that every event involved in the act is caused by some other event; and we must not say that the act is something that is not caused at all. The possibility that remains, therefore, is this: We should say that at least one of the events that are involved in the act is caused, *not* by any other events, but by something else instead. And this something else can only be the agent – the man.[408]

The agent causation theory of free actions has at least two advantages. It avoids the theoretical problems of compatibilism and hard determinism, and it accords well with the ordinary thinking about reasons why people are responsible for their behavior. We tend to think that an action must be produced by an *agent* in some deeper sense than simply by his visible body as a necessary condition for deserving praise or blame for the action. Moreover, references to a person's desires, values and beliefs are rarely seen as providing an *excuse* for his immoral behavior. One should still be held accountable for raping a victim despite the fact that he *wanted* to rape her very much or believed that it was the best thing to do. The implication of this intuition is that our inner mental states are in some sense under our own control, or at least should be under our control in most instances of willful action. Chisholm builds on Leibniz's insight when he argues that "our desires may incline without necessitating,"[409] thus leaving a room for control by the agent.

The claim that we are able at least *to control* the strength of our desires (if not to choose our desires in the first place) is further confirmed by a number of empirical studies. Alfred Mele cites the following evidence:

---

[408] Chisholm, 1992, p. 147.

[409] *Ibid*., p. 153.

> The studies of psychologists George Ainslie and Walter Mischel indicate that desire-strength is subject to intentional control. […] If behavioral therapists and our own experience may be trusted, we are often able to bring the strength of our desires into line with our better judgments, unless a desire is *irresistible*.[410]

That psychopaths are *not* normally operating under the strength of irresistible desires and impulses is admitted even by those who want to absolve them from responsibility.[411] It follows that even the psychopaths *could* have chosen differently on a given occasion had they had different goals, where the possibility expressed by this subjunctive is of a more intimate and more *real* kind than the one functioning in the usual 'possible worlds' discourse.

The ability to exercise such control does not involve the presence of an alternative desire of a greater intensity which wins out in the competition. The control must be in the hands of the free agent himself rather than any of his mental states. This, Chisholm concludes, leads to the following view of a responsible agent:

> If we are responsible, and if what I have been trying to say is true, then we have a prerogative which some would attribute only to God: each of us, when we act, is a prime mover unmoved. In doing what we do, we cause certain events to happen, and nothing – or no one – causes us to cause those events to happen.[412]

Among the events and states we are able to cause will undoubtedly be the ones regulating the adoption of goals and establishing the priority of goals and values that are already adopted. Only this scenario, it seems, would allow us to preserve the intuitive sense of freedom necessary for moral responsibility. The agent is not a passive observer of the battle between the opposing desires that rages 'in his mind', but rather the active originator of these desires

---

[410] Mele, 1992, pp. 80-81.

[411] Cf. A. Duff: "[A psychopath] cannot control his actions in the light of any rational concerns or values, *not* because his impulses are strictly irresistible, but because he has no conception of rational values as providing reasons for actions." (1977, p. 199).

[412] Chisholm, 1992, p. 152.

(at least of some of them) and the controlling center that is capable of determining their respective intensity.

The conception of free action sketched above has an immediate but radical corollary: the question why a free agent has chosen to act on this particular value rather than on a different one *cannot* be fully answered. No matter how specific we are in describing the person's inner mental states at the moment of decision, there will always be a residuum that defies the explanatory efforts - not because of lack of relevant information but on principled grounds. This important skeptical implication of Chisholm's view is explicitly acknowledged in the following passage:

> No set of statements about a man's desires, beliefs, and stimulus situation at any time implies any statement telling us what the man will try, set out, or undertake at that time. This means that there can be no complete science of man. If we think of science as a matter of finding out what laws happen to hold, and if the statement of a law tells us what kinds of events are caused by what other kinds of events, then there will be human actions which we cannot explain by subsuming them under any laws.[413]

A person's beliefs and desires can obviously incline him in the direction of a particular action. But if an action is free in the sense explained, there must be a variable that is not captured by any descriptions of a person's mental states, but the one that may frustrate the original inclination, however strong it is (short of being irresistible). The variable in question is the freely adopted *goal* of action. Richard Taylor, who largely shares Chisholm's view on freedom and responsibility, echoes the same thought when he writes: "If human behavior is

[413] *Ibid*.

ever purposeful, as in fact it usually is, then it can never be understood in terms of the concepts of physical science."[414]

Free agents, on this view, may be unique in the sense that, unlike the case with all other occurrences in the universe which can and should be explained in terms of their efficient causes, the intentional actions of human beings essentially contain a reference to final goals of action.[415] Indeed, reference to a final goal is *constitutive* of a genuine action (as opposed to a reflexive motion of limbs), and hence, "no description or analysis of [purposeful behavior] in terms of the concepts of physical science will work."[416] The desired end is what ultimately explains the occurrence of an action, but the adoption of this end itself remains a free and, in the final analysis, *unexplainable* choice of a responsible agent.

This view essentially refuses to follow Hobbes and other compatibilists who consider an action as a kind of *effect* of antecedent desires. Rather, the defended model of intentional action here is the one that sees *goal-directedness* as its distinctive feature. And a decision to adopt a certain goal or to set the level of its priority is a product of practical rationality. As Thomas Pink writes, "[an action] is a mode of exercising reason; what distinguishes action is not a special kind of cause, but a special kind of rationality."[417] The capacity to set goals for oneself and to pursue them is also what Kant sees as being the most valuable capacity of rational beings. Free action involves a free decision to adopt a certain end, which involves a free exercise of practical reasoning. If freedom of will and responsibility should be

---

[414] Taylor, 1966, p. 224.

[415] We can agree with Jean Hampton (1996) that any theory that makes references to final goals is 'unscientific' without yet sharing his intuition that this unscientific status must be an inherently bad thing.

[416] *Ibid*., p. 225.

[417] Pink, 2004, p. 98.

preserved, Kant argues, our goals cannot be imposed on us by external environment or fully determined by the force of existing motivations:

> A *goal* is an *object* of free choice, the representation of which determines it to an action (by which the object is brought about). Every action, therefore, has its goal; and since no one can have a goal without *himself* making the object of his choice into a goal, to have any goal of action whatsoever is an act of *freedom* on the part of acting subject, not an effect of *nature*.[418]

Kant's position implies that the initial decision of goal-adoption is itself a kind of free action, which involves practical reasoning and whose rationality can be appraised not only in terms of its efficiency to achieve some further goal, but also with regard to the object of choice being indeed *choiceworthy*. How exactly this decision comes to pass, however, is something that must, according to Kant, forever remain "theoretically unknowable". An account of purposeful human behavior that fits best with common intuitions and our other fundamental interests and concerns is achievable only at the expense of limiting the scope of possible theoretical knowledge about the causes of one's free actions. But this epistemological price is definitely worth paying for the resulting metaphysical picture of human agency.

## 3. Summary and Concluding Remarks

I have argued in the preceding sections that amoralists are morally responsible for their actions, that responsibility requires freedom of will, and that freedom of will involves freedom of goal adoption. Why is it, then, that some people routinely choose to follow the immoral values, or else, downgrade the status of the ends of morality to the level where an action in conformity with moral requirements is possible only in rare cases where no other

[418] Kant, I., 1996, p. 516.

opposing desires intervene? I believe that in the final analysis this 'why-question' *cannot* have an answer.

If the preceding reasoning is sound, and amoralism as a state of character is not something that *happens* to a human being but is rather something that he willfully chooses, then any explanation of a person's behavior must come to a forced stop once it reaches the realm of free decision. Most humans have an ability to resist the coercion of desires and impulses, and, hence, their behavior cannot be fully explained by reference to these internal motivating causes. Even if all the relevant desires and beliefs preceding an immoral action are specified, it will still remain unexplained why a person chose to follow these desires rather than to resists them in view of some other goal or value. Indeed, if Kant's insight in the "Religion Within the Boundaries of Mere Reason" and our common intuition can be trusted, the free choice must remain essentially unexplain*able*:

> Whenever we therefore say, "The human being is by nature good," or, "He is by nature evil," this only means that he holds within himself a first ground (to us inscrutable [*uns unerforschlichen*]) for the adoption of good or evil maxims, and that he holds this ground *qua* human, universally.[419]

The ground that is 'inscrutable to us' is the realm of freedom which resists rational understanding in terms of determining physical or psychological causes. If no causal factor necessitates the agent to choose action A over B, or to give higher preference to a certain value, then the traditional explanatory approach that proceeds by explicating the causes of events will always be inadequate when applied to human choices. This is not to say that referring to various contributing factors to amoral behavior is always inappropriate. Martha Stout, for instance, cites the heritability studies that suggest that "a person's tendency to

[419] Kant, 1998, p. 47; 6:21.

possess certain sociopathic characteristics is partially borne in the blood, perhaps as much as 50 percent so."[420] The genetic predisposition to amoral states of character may well be one of such contributing factors whose presence makes a statistical difference in predicting the cases of anti-social behavior. Yet, unless the genetic abnormality in question leads to *irresistible* desires, the prediction will never be absolutely certain. Benjamin Wolman expresses a more reasonable attitude when he honestly admits that at present "there is no agreement as to what causes sociopathy or what are the dynamics of sociopathic behavior."[421] He arrived at this skeptical conclusion after carefully examining the existing empirical data on psychopaths, but we have seen that the lack of agreement among the researchers on the causes of psychopathy can in turn be explained by the peculiarity of the subject matter they are dealing with, namely, a free human agency.

The above insights have a quite general application. Morally praiseworthy behavior is equally unexplainable once we press the 'why' question far enough, and this, in turn, ensures the appropriateness of praise we bestow on a person. The assent one gives to a certain principle of action cannot be fully accounted for by any of the usual causal factors, such as heredity, upbringing or internal psychological states. The assent stems from the *person* himself, and that is precisely what allows us to blame or praise the *person* rather than his beliefs, desires or external surroundings. No matter how appealing it is to cite the external factors in explaining one's behavior, our natural desire for intellectual closure must be given up if we want to preserve the acceptable sense of moral responsibility as applied to human agents.

---

[420] Stout, M., 1995, p. 122.

[421] Wolman, B., 1987, p. 87.

I shall conclude this chapter by a quote from Richard Bernstein, who builds on Kant's insight cited above, and nicely summarizes the main idea of this section:

> In the final analysis, we cannot explain why one person chooses to become good and another chooses to become evil. To imagine that we could explain this would be in effect to deny that our will is radically free. So, far from its being some sort of deficiency, it is Kant's way of acknowledging a profound moral truth about our radical free choice. Human beings are responsible for the choices they make, but *ultimately*, we cannot explain why they make the moral choices they do; we cannot explain the "ultimate subjective ground of the adoption of moral maxims" – whether for good or for evil. Not only is this inscrutable; it *must* be inscrutable, because this is what it means to be a free and responsible person.[422]

The domain of freedom is one aspect of human existence where, at a certain level of questioning, intellectual resignation is entirely appropriate. It is not a premature despair of a person faced with a formidable subject matter, but an attitude born out of a fuller appreciation of the nature of free choice and the limits of our understanding. The investigation into the roots of amoralism undertaken in this work with its largely negative results is meant, among other things, to illustrate the truth of that insight.

*[Jan. 2005 – Oct. 2007]*

[422] Bernstein, 2002, p. 45.

# BIBLIOGRAPHY


- American Psychiatric Association *Diagnostic and Statistical Manual for Mental Disorders*, 4th ed., (DSM IV), Washington, DC: 1994.
- Arendt, Hannah. *Eichmann in Jerusalem: A Report on the Banality of Evil*. New York: The Viking Press, 1963.
- Aristotle. *The Basic Works of Aristotle*, ed. by Richard McKeon, New York: Random House, 1941.
- Arrington, Robert L. "Practical Reason, Responsibility and the Psychopath." *Journal for the Theory of Social Behavior*, vol. 9, No. 1 (1979):71-89.
- Audi, Robert. *Practical Reasoning and Ethical Decision*. London and New York: Routledge, 2006.
- Ayer, A. J. *Language, Truth and Logic* (1936). New York Dover Publications, 1952.
- Baier, Kurt. *The Moral Point of View*. New York: Random House, 1965.
- Baron-Cohen, S., Ring, H. A., et al. "The Amygdala Theory of Autism." *Neuroscience Biobehavior Review*, vol. 24, (2000): 355-64.
- Benedict, Ruth. *Patterns of Culture*. (1935). Boston: Houghton Mifflin, 1989.
- Bennett, Jonathan. "The Conscience of Huckleberry Finn." *Philosophy*, vol. 49, no. 188, (1974): 123-34.
- Bernstein, Richard J. *Radical Evil: A Philosophical Investigation*. Cambridge: Polity Press, 2002.
- Blair, R. J. R. "A Cognitive Developmental Approach to Morality: Investigating the Psychopath." *Cognition*, vol. 57, (1995): 1-29.
- _____. "The Emergence of Psychopathy." *Cognition*, vol. 101 (2006): 414-42.

- Blum, Lawrence A. *Friendship, Altruism, and Morality*. London: Routledge & Kegan Paul, 1980.
- Bower, Bruce. "The Predator's Gaze." *Science News*, vol. 170, 2006: 371-83.
- Boyd, Richard. "How to be a Moral Realist." *Essays on Moral Realism*. Geoffrey Sayre-McCord, ed., Ithaca, NY: Cornell University Press, 1988: 182-217.
- Brink, David. "Moral Motivation." *Ethics*, vol. 108, (1997): 4-32.
- Cahn, Steven M. "The Happy Immoralist." *Journal of Social Philosophy*, vol. 35, (2004): 1-1.
- Cherry, Christopher. "Agreement, Objectivity and the Sentiment of Humanity." *Nature and Conduct*. R. S. Peters, ed., New York: St. Martin's Press, 1975: 83-98.
- Chisholm, Roderick. "Human Freedom and the Self." *Free Will.* D. Pereboom, ed., Cambridge: Hackett Publishing Company, Inc., 1997:143-55.
- Chomsky, Noam. *Language and Problems of Knowledge: The Managua Lectures*. Cambridge: MIT Press, 1988.
- Cleckley, Hervey. *The Mask of Sanity*. New York: The New American Library, 1982.
- Conolly, John. *An Inquiry Concerning the Indications of Insanity* (1830). London: Dawsons, 1964.
- Copp, David. "Belief, Reason, and Motivation." *Ethics*, vol. 108, (1997): 33-54.
- Damasio, A. R. *Descartes' Error: Emotions, Reason, and the Human Brain*. New York: Grosset/Putnam, 1994.
- Dancy, Jonathan. *Moral Reasons.* Oxford: Blackwell Publishers, 1993.
- Davis, Nancy and Jamieson, Dale. "Ross on the Possibility of Moral Theory." *The Journal of Value Inquiry*, vol. 21, (1987): 225-34.

- Deigh, John. "Empathy and Universalizability." *Ethics*, vol. 105, no. 4, (1995): 743-63.
- Dennett, Daniel. *Consciousness Explained*. Boston: Little, Brown and Col., 1991.
- Devitt, Michael. "Moral Realism: A Naturalistic Perspective." *Croatian Journal of Philosophy*, vol. 2, (2002): 1-15.
- Dostoyevsky, Fyodor. *The Possessed* (1872). trans. by C. Garnett. New York: The Modern Library, 1936.
- _____. *Notes from the Underground* (1864). trans. by Jessie Coulson. Penguin Books, 1988.
- Duff, Antony. "Psychopathy and Moral Understanding." *American Philosophical Quarterly*, vol. 14, Number 3, (July 1977): 189-200.
- Ekman, Paul. "An Argument for Basic Emotions." *Cognition and Emotion*, vol. 6, (1992): 169-200.
- Ekman, P. and Friesen, W. V. "Constants Across Cultures in the Face and Emotion." *Journal of Personality and Social Psychology,* v. 17 (2), (1971): 124-29.
- _____. "A New Pan-cultural Facial Expression of Emotion." *Motivation and Emotion*, 10 (2), (1986): 159-68.
- Eisenberg, Nancy. "The Development of Empathy-Related Responses." *Moral Motivation through the Life Span*, Nebraska Symposium on Motivation, vol. 51, (2005): 73-117.
- Evans, Dylan. *Emotion: The Science of Sentiment*. Oxford: Oxford University Press, 2001.
- Flack, J. and Frans de Waal. "Any Animal Whatever." *Evolutionary Origins of Morality*., Leonard Katx, ed., Bowling Green, 2000.

- Flew, Antony. *Crime or Disease?* London: Macmillan, 1973.
- Foot, Philippa. "Morality as a System of Hypothetical Imperatives." (1971). *Virtues and Vices*, Oxford: Blackwell Publishers, 1978: 157-63.
- _____. "Moral Beliefs." (1959). *Virtues and Vices*, Oxford: Blackwell Publishers, 1978: 110-31.
- _____. "Reasons for Action and Desires." (1972). *Virtues and Vices*, Oxford: Blackwell Publishers, 1978: 148-56.
- _____. "Moral Arguments." *Mind*, vol. LXVII, (1958): 502-13.
- Fowles, D., Sutker, P., et al., eds. *Special Focus on Psychopathy and Antisocial Personality: A Developmental Approach.* New York: Springer Publisher, 1994.
- Frankena, William K. "Obligation and Motivation in Recent Moral Philosophy." *Essays in Moral Philosophy*. A. I. Melden, ed., Seattle: University of Washington Press, 1958: 40-81.
- _____. *Ethics*. Upper Saddle River, New Jersey: Prentice Hall, 1963.
- Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy*, vol. 68, no. 1, (1971): 5-20.
- Frith, Uta. *Autism: Explaining the Enigma*. Oxford: Blackwell, 1989.
- _____. ed., *Autism and Asperger Syndrome*. New York: Cambridge University Press, 1991.
- Fromm, Erich. *The Anatomy of Human Destructiveness*. London: Penguin, 1977.
- Garner, Richard. *Beyond Morality*. Temple University Press, Philadelphia, 1994.
- Gert, Bernard. *Morality: Its Nature and Justification*. Oxford University Press, New

York, 1998.

- Giannangelo, Stephen J. *The Psychopathology of Serial Murder: A Theory of Violence*. London: Praeger series, 1996.
- Glannon, Walter. "Psychopathy and Responsibility." *Journal of Applied Philosophy*. Vol. 14, No. 3 (1997): 263-75.
- Glover, Jonathan. *Humanity: A Moral History of the Twentieth Century*. New Haven and London: Yale University Press, 1999.
- _____. *Responsibility*. Routledge, New York and London: Humanity Press, 1970.
- Goddard, H. *The Criminal Imbecile*. New York: Macmillan, 1915.
- Grant, C. M., Boucher J., et al. "Moral Understanding in Children with Autism." *Autism*, vol. 9, (2005): 317-31.
- Griffiths, Paul E., *What Emotions Really Are*. Chicago: The University of Chicago Press, 1997.
- Greene, Joshua. "Cognitive Neuroscience and the Structure of the Moral Mind." *The Innate Mind: Structure and Contents*. S. Laurence, P. Carruthers, and S. Stich, eds., New York: Oxford University Press, 2005.
- Greenspan, P. *Emotions and Reasons*. New York: Routledge, 1988.
- Haksar, Vinit. "Aristotle and the Punishment of Psychopaths." *Philosophy*, vol. 39, (1964): 323-40.
- _____. "The Responsibility of Psychopaths." *The Philosophical Quarterly,* vol. 15, No 59, (1965): 135-45.
- Hampton, Jean. "The Nature of Immorality." *Social Philosophy and Policy*, vol. 7,

no. 1, (1989): 115-34.

- _____. *The Authority of Reason*. Cambridge: Cambridge University Press, 1996.
- Hare, Richard M. *The Language of Morals*. Oxford: Oxford University Press, 1952.
- _____. *Freedom and Reason*. Oxford: Oxford University Press, 1963.
- _____. *Moral Thinking.* Oxford: Clarendon Press of Oxford University Press, 1981.
- Hare, Robert D. *Without Conscience: the Disturbing World of the Psychopaths among Us*. New York: Pocket Books, 1993.
- Harman, Gilbert. "Is There a Single True Morality?" in David Copp & D. Zimmerman, eds., *Morality, Reason and Truth*. Rowman & Allanheld, 1985: 27-48.
- _____. *The Nature of Morality: An Introduction to Ethics*. New York: Oxford University Press, 1977.
- Harrington, Alan. *Psychopaths*. New York: Simon and Schuster, 1972.
- Herve, H., Yuille J., eds. *The Psychopath: Theory, Research, and Practice*. Lawrence Erlbaum Associates, 2007.
- Hobbes, Thomas. *Leviathan* (1651). ed. Edwin Curley, Indianapolis: Hackett, 1984.
- Hume, David. *A Treatise on Human Nature*. (1739), ed. by L. A. Selby-Bigge, 2nd ed., revised by P. H. Nidditch. Oxford: Clarendon Press, 1978.
- _____. *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals.* (1777), ed. by L. A. Selby-Bigge, Oxford: Clarendon Press, 2nd ed., 1966.
- Jaffa, Harry V. *Homosexuality and the Natural Law*. Claremont, CA: The Claremont

Institute of the Study of Statesmanship and Political Philosophy, 1990.

- Jaspers, Karl. *General Psychopathology*. Hoenig and Hamilton, trans., Chicago: University of Chicago Press, 1963.
- Jenkins, Richard L. "The Psychopathic or Antisocial Personality." *Journal of Nervous and Mental Disease*, vol. 131, (1960): 318-34.
- Joyce, Richard. *The Myth of Morality.* Cambridge: Cambridge University Press, 2002.
- Kamm F. M. "Rescuing Ivan Ilych: How We Live and How We Die." *Ethics*, vol. 113, (2003): 202-34.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals* (1785). Mary Gregor, trans. & ed.; with intro. by Christine M. Korsgaard. Cambridge: Cambridge University Press, 2001.
- _____. *Religion Within the Boundaries of Mere Reason* (1793). Allen Wood and G. Giovanni, trans. & eds., Cambridge: Cambridge University Press, 1998.
- _____. *The Doctrine of Virtue* (1795). Mary J. Gregor, trans. & ed., New York: Harper & Row, 1964.
- _____. "What is Enlightenment?" *Practical Philosophy: The Cambridge Edition of the Works of Immanuel Kant*. Cambridge: Cambridge University Press, 1996.
- _____ ."Metaphysics of Morals" (1795). *Practical Philosophy: The Cambridge Edition of the Works of Immanuel Kant*. Cambridge: Cambridge University Press, 1996.
- Kennett, Jeanette. "Autism, Empathy and Moral Agency." *The Philosophical Quarterly*, vol. 52, No. 208, (July 2002): 340-57.

- Kohlberg, Lawrence. "Stage and Sequence: The Cognitive Developmental Approach to Socialization." *Handbook of Socialization Theory and Research*, D. A. Goslin, ed., New York: Rand McNally, 1969: 325-480.
- _____. "From Is to Ought: How to Commit the Naturalistic Fallacy and Get Away With It." *Cognitive Development and Epistemology*, T. Mischel, ed., New York: Academic Press, 1971: 151-235.
- _____. "The Claim to Moral Adequacy of a Highest Stage of Moral Judgment." *Journal of Philosophy*, vol. 70, No. 18, (1973): 630-46.
- _____. "Moral Stage and Moralization." *Moral Development and Behavior*. T. Lickona, ed., New York: Holt & Winston, 1976: 84-107.
- Lapsley, Daniel K. *Moral Psychology*. Oxford: Westview Press, 1996.
- Little, Margaret Olivia. "Seeing and Caring: The Role of Affect in Feminist Moral Epistemology." *Hypatia*, 10:3, 1995: 117-37.
- Lykken, David T. *The Antisocial Personalities*. New Jersey: Lawrence Erlbaum Associates Publishers, 1995.
- Mackie, John L. *Ethics: Inventing Right and Wrong*. Penguin, 1977.
- McGinn, Colin. *Ethics, Evil, and Fiction*. Oxford: Clarendon Press, 1997.
- McNaughton, David. *Moral Vision*. New York: Blackwell, 1988.
- Mele, Alfred R. *Springs of Action: Understanding Intentional Behavior*. New York: Oxford University Press, 1992.
- Michaud, Stephen, & Aynesworth, Hugh. *Ted Bundy: Conversations with a Killer*. Barnes & Noble, Inc., 2005.
- Mill, J. S. *Utilitarianism*. (1861). Roger Crisp ed., Oxford: Oxford University Press,

1998.

- Milo, Ronald D. *Immorality*. Princeton: Princeton University Press, 1984.
- Monro, David. H. *Empiricism and Ethics*. Cambridge: Cambridge University Press, 1967.
- Moore, George E. *Principia Ethica* (1903). Cambridge: Cambridge University Press, 1962.
- Murphy, Jane M. "Psychiatric Labeling in Cross-Cultural Perspective." *Science*, vol. 191, (1976): 1019-28.
- Murphy, Jeffrie G. "Moral Death: a Kantian Essay on Psychopathy." *Ethics*, vol. 82, (1972): 284-98.
- Nagel, Thomas. *The Possibility of Altruism*. Oxford: Clarendon Press, 1970.
- Novell-Smith, P. H. *Ethics*. Baltimore: Penguin Books, 1954.
- Nielsen, Kai. "Must the Immoralist Act Contrary to Reason?" in David Copp & D. Zimmerman, eds., *Morality, Reason and Truth*. Rowman & Allanheld, 1985: 212-27.
- _____. *Why Be Moral?* Buffalo: Prometheus Books, 1989.
- Nietzsche, Friedrich *Thus Spake Zarathustra*. New York: The Modern Library, 1950.
- Piaget, Jean. *The Moral Judgment of the Child* (1932). New York: Free Press, 1965.
- Pinel, Philippe. *A Treatise on Insanity* (1801). Tr. by D. D. Davis, Washington: University Publication of America, 1977.
- Pink, Thomas. *Free Will*. Oxford: Oxford University Press, 2004.
- Pizarro, David. "Nothing More than Feelings? The Role of Emotions in Moral Judgment." *Journal for the Theory of Social Behavior*, vol. 30, No. 4, (2000): 355-75.
- Plato. *The Collected Dialogues.* ed. by Edith Hamilton and Huntington Cairns.

Princeton University Press, 1996.

- Pojman, Louis P. *Ethics: Discovering Right and Wrong*. Thomson Wadsworth, 5th edition, 2006.
- Preston, Stephanie and Frans de Waal. "Empathy: Its Ultimate and Proximate Bases." *Behavioral and Brain Sciences*, vol. 25, (2002): 35-58.
- Prichard, James Cowles. *A Treatise on Insanity and Other Disorders Affecting the Mind* (1835). New York: Arno Press, 1973.
- Pritchard, Michael S. "Responsibility, Understanding, and Psychopathology," *The Monist*, vol. 58, no. 4, (1974): 630-45.
- Prinz, Jesse. "Imitation and Moral Development." *Perspectives on Imitation: From Cognitive Neuroscience to Social Science.* Susan Hurley and Nick Charter, eds. Vol. 2: *Imitation, Human Development, and Culture*. MA: MIT Press, 2005: Chapter 25.
- Quinton, Anthony. "Has Man an Essence?" *Nature and Conduct*. R. S. Peters, ed., New York: St. Martin's Press, 1975: 8-34.
- Rachels, James. *The Elements of Moral Philosophy*. McGrawl-Hill, 3rd ed., 2000.
- Rand, Ayn. *The Virtue of Selfishness*. New York: New American Library, 1965.
- Rawls, John. *A Theory of Justice*. Cambridge: Belknap Press of Cambridge University Press, 1971.
- _____. "Kantian Constructivism in Moral Theory." Darwall, Gibbard, Railton, eds., *Moral Discourse & Practice*, New York: Oxford University Press, 1997: 247-66.
- Reid, Thomas. *Essays on the Intellectual Powers of Man*. Cambridge: MIT Press, 1969.

- Ross, Steven L. "Evaluating the Emotions." *The Journal of Philosophy*, vol. 81, no. 6, 1984: 309-26.
- Ross, William D. *The Right and the Good.* Oxford: Clarendon Press, 1930.
- Rousseau, Jean-J. *Emile* (1760). London: Everyman's Library, 1974.
- Rush, Benjamin. *Medical Inquiries and Observations Upon the Diseases of the Mind* (1811). New York: Library of New York Academy of Medicine, 1962.
- Ryle, Gilbert. "Conscience and Moral Convictions." *Analysis*, vol. vii, (1940): 3-16.
- _____. "On Forgetting the Difference Between Right and Wrong." *Essays in Moral Philosophy*. ed. A. I. Melden. Seattle: University of Washington Press, 1958: 147-59.
- Sacks, Oliver W. *An Anthropologist on Mars*. New York: Vintage Books, 1995.
- Samenow, E. Stanton. *Inside the Criminal Mind*. New York: Times Books, 1984.
- Sartre, Jean-Paul. "Existentialism." (1945) in *Existentialism and Human Emotions*, tr. Bernard Frechtman. New York: Philosophical Library, 1957: 9-40.
- _____. *Sketch for a Theory of the Emotions*. London: Methuen, 1971.
- Scanlon, Thomas M. *What We Owe to Each Other*. Cambridge and London: The Belknap Press of Harvard University Press, 1998.
- Schafer-Landau, Russ. *Moral Realism: A Defense*. Oxford: Oxford University Press, 2003.
- Schalling, D., Rosen A., "Porteus Maze Differences Between Psychopathic and Non-Psychopathic Criminals." *British Journal of Social Clinical Psychology*, vol. 7, (1968): 224-8.
- Scheler, Max. *The Nature of Sympathy*. Trans. by Peter Heath. London: Routledge &

Kegan Paul, 1954.

- Schopenhauer, Arthur. *On the Basis of Morality*. (1841) Trans. E. F. Payne. Oxford: Berhahn Books, 1995.
- Simpson, Peter. *Goodness and Nature: A Defense of Ethical Naturalism*. Martinus Nijhoff Publishers, 1987.
- Skillen, Anthony. "Conflicting Strains in Marxian Anti-Moralism." *Marx and Morality*, K. Nielsen, S. Patten eds., Ontario: Canadian Association for Publishing in Philosophy, 1981: 155-70.
- Slote, M. *Moral from Motives*. Oxford: Oxford University Press, 2001.
- Smith, Robert J. "The Psychopath as Moral Agent." *Philosophy and Phenomenological Research*. Vol. xiv, No. 2, (1984): 177-93.
- Smith, Michael A. *The Moral Problem*. Oxford: Basil Blackwell Publishers, 1994.
- Solomon, Robert C. *On Ethics and Living Well*. Thomson Wadsworth, 2006.
- _____. *The Passions*. New York: Anchor, 1977.
- Strawson, Peter. *Freedom and Resentment and Other Essays.* London: Methuen, 1974.
- Stocker, Michael. "Desiring the Bad." *Journal of Philosophy*, vol. 76, (1979): 738-53.
- Stout, Martha. *The Sociopath Next Door*. New York: Broadway Books, 2005.
- Sutker, P. B., Moan, C., *et al*. "Assessment of Cognitive Control in Psychopathic and Normal Prisoners." *Journal of Psychopathology and Behavioral Assessment*. Vol. 5, (Dec. 1983): 275-87.
- Taylor, Richard. *Action and Purpose*. Englewood Cliffs: Prentice Hall, 1966.
- Timmons, Mark C. *Moral Theory: An Introduction.* New York: Rowman & Littlefield Publishers, 2002.

- Tranel, Daniel. "Acquired Sociopathy." *Progress in Experimental Personality and Psychopathology Research*. D. Fowles, P.Sutker, eds., New York: Springer Publishing Company, 1994: 285-312.
- Turnbull, Colin. *The Mountain People*. Touchstone, 1987.
- van Inwagen, Peter. *An Essay on Free Will*. Oxford: Clarendon press, 1983.
- Vetlesen, Arne Johan. *Perception, Empathy, and Judgment: An Inquiry into the Preconditions of Moral Performance*. The Pennsylvania State University Press, 1994.
- Velleman, David J. "The Possibility of Practical Reason." *Ethics*, vol. 106, (July 1996): 694-726.
- Vitrano, Christine. "Happiness and Morality." *Happiness*. Cahn, S., Vitrano, C., eds., Oxford University Press, 2008: 267-72.
- Watson, Gary. "Responsibility and the Limits of Evil." *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Cambridge: Cambridge University Press, 1987: 256-86.
- Wilde, Oscar. *The Picture of Dorian Gray* (1890). New York: Dolphin Books, 1960.
- Williams, Bernard. *Problems of the Self.* Cambridge: Cambridge University Press, 1973.
- _____. *Morality: An Introduction to Ethics*. New York: Harper & Row, 1972.
- _____. *Moral Luck*. Cambridge: Cambridge University Press, 1981.
- _____. *Ethics and the Limits of Philosophy*. Cambridge: Harvard University Press, 1985.
- Wilson, Catherine. *Moral Animals: Ideals and Constraints in Moral Theory*. New York: Clarendon Press, 2004.

- Wolf, Susan. "Moral Saints." *The Journal of Philosophy*, vol. 79, No. 8, (1982): 419-39.
- Wolman, Benjamin B. *The Sociopathic Personality*. New York: Brunner/Mazel, 1987.
- _____. *Antisocial Behavior: Personality Disorders from Hostility to Homicide*. Amherst, New York: Prometheus Books, 1999.
- Zahn-Waxler, C., Radke-Yarrow, M., et al. "Development of Concern for Others." *Developmental Psychology*, vol. 28, (1992): 126-36.
- Zajonc, R. B. "Feeling and Thinking." *American Psychologist*, vol. 35, (1980): 151-75.