Peter Lucas

Ethics and Self-Knowledge

Respect for Self-Interpreting Agents

Springer

ETHICS AND SELF-KNOWLEDGE

LIBRARY OF ETHICS AND APPLIED PHILOSOPHY

VOLUME 26

For further volumes:
http://www.springer.com/series/6230

ETHICS AND SELF-KNOWLEDGE

Respect for Self-Interpreting Agents

by

PETER LUCAS
University of Central Lancashire, UK

Springer

Dr. Peter Lucas
School of Education and Social Sciences
University of Central Lancashire
Preston
Lancashire PR1 2HE
UK
plucas1@uclan.ac.uk

The Other by rising up confers on the for-itself a being-in-itself-in-the-midst-of-the-world as a thing among things. This petrifaction in in-itself by the Other's look is the profound meaning of the myth of Medusa

Jean-Paul Sartre, *Being and Nothingness*

It is not always from pride that a mistress refuses to yield to the caprices of her lover: she would fain have to do with an adult who is living out a real moment of his life, and not with a little boy telling himself stories

Simone de Beauvoir, *The Second Sex*

For Michael Hammond – who showed me how, and why

Acknowledgments

All philosophy begins in dialogue; and in the case of this book, which has had a lengthy gestation, the relevant dialogues have been many and varied. It would not be possible to mention by name all those who have contributed, in one way or another. Nevertheless a few names stand out. Firstly, the staff of Lancaster University Philosophy Department, with whom I took my first steps in the discipline. In particular: Michael Hammond, Jane Howarth and Alan Holland. It was their teaching, and their example, that first enabled me to think of myself as someone who might one day write a book like this. An equally long-standing debt is to David Littlewood, who in endless exhilarating conversations showed a surer touch in ethics than I will ever have. I am also indebted to current and former colleagues at the University of Central Lancashire: to Peter Herissone-Kelly, who endured numerous interruptions with only minor irritation, and served as a sounding board in the development of some of the central ideas; to Niall Scott, who drew my attention to Martha Nussbaum's work on sexual objectification; to Johan Siebers, who provided encouragement in the darker hours; and to Matti Häyry, who arranged invaluable relief from teaching at a crucial stage, and also read and commented on the entire manuscript. I would also like to thank the anonymous reviewer for Springer, for detailed and very helpful comments, and for prompting me to use the terminology of *interpretive* moral wrongs. The staff of the Department of Humanities at the University of Bolton kindly invited me to present some of my ideas at their 2008 Philosophy Seminar series, and their comments had a tangible impact on the finished work. Finally, I must record my gratitude to all my students, past and present, for (as the saying goes) 'keeping me honest'.

Contents

Chapter 1
Introduction

Abstract The charge that an individual or group is guilty of *objectifying* some other individual or group, is, or once was, a fairly familiar feature of ethical and political discussion. The 'objectification' charge combines elements that established ethical and political theories seem content to consider separately, if they consider them at all. Objectification concerns the way we treat others, the way we understand them, and the way we represent them. This Introduction outlines the aims and structure of the book, which sets out to subject the concept of objectification to a form of critical scrutiny it has not previously received. In considering why contemporary applied and professional ethics has not previously paid close attention to objectification, it uncovers an apparent crisis afflicting the ethical principle of respect for persons. This in turn leads to the identification of a distinct class of moral wrongs – which I term *interpretive* moral wrongs – to which objectification belongs, and which have not previously been explicitly distinguished.

1.1 The Puzzle of Objectification

The charge that an individual or group is guilty of objectifying some other individual or group, is, or once was, a fairly familiar feature of ethical and political discussion. The term 'objectification' has a superficially ethically-neutral sense in the Hegelian philosophical tradition, where it refers to the productive processes by which consciousness externalises itself, and renders itself objective in its own creations (a process which is held to first make certain forms of self-consciousness possible).[1] The term is also used in an entirely normatively neutral sense in analytic meta-ethics.[2] In its most familiar sense though, it refers, somewhat loosely, to the treatment of a person as if they were an object, or to their reduction to the status of an object.[3]

I first became aware of this concept of objectification in the early nineteen-eighties, when I encountered the claim, promulgated by what was at that time an

------

[1] Marx 1967, especially pp. 315–37.

[2] Mackie 1990, pp. 42–3.

[3] In this sense the term probably derives from Sartre's concept of 'objectivation' – see e.g. Sartre 1989, p. 161. Hazel Barnes glosses objectivation in this sense as the process of 'making an object out of something or someone' (Sartre 1989, p. 633).

energetic British feminist movement, that our male-dominated culture practices a very general objectification of women. The primary targets of the charge were the pornography, advertising and entertainment industries, and the news media (profitably combined in some sectors of the press). These were attacked both on the basis that the ways in which they depicted women were objectifying in and of themselves (treating them as objects, or reducing them to the status of objects), and on the basis that they served to foster objectifying attitudes and behaviours in the population as a whole.

Several features of this concept of objectification caught my attention. Firstly, there was the thought that something more all-embracing than a political ideology could be invoked in order to explain and critique everyday social behaviour. To be guilty of objectifying women, it did not seem necessary to be a card-carrying patriarchalist, and even the best-intentioned male had reason to question his own habitual attitudes and responses. Secondly, there was the thought that, while the existence of a general male tendency to objectify women might serve to explain an individual's treatment of others, objectification is not apparently reducible to overt behaviour. One could be guilty of objectifying another without visibly *doing* anything to them. (Objectification, thus understood, seemed almost to have the character of a 'thoughtcrime'.) Thirdly, the objectification charge seemed to combine – in a manner that still strikes me as highly intriguing – elements that traditional ethical and political theories seem content to consider separately (if they consider them at all). Objectification concerned the way we treat others, the way in which we know and understand them, and the way we represent them. Even, it seemed, the way we see them. Thus the idea of objectification interested me, and continues to interest me, because it registers the fact that there are important reciprocal connections between ethics, political authority, and epistemic authority, which ought to command the attention of anyone seriously concerned with the avoidance of ethical and political abuses.

The final fascinating feature of the concept of objectification was that nobody seemed to know precisely what it involved, or how it worked. This latter feature has become increasingly evident during the intervening decades – in which many of the fundamental political goals of seventies and eighties feminism have supposedly been achieved, but which have simultaneously seen a proliferation of the kinds of commercial and media portrayals of women that were the prime targets of the original anti-objectification campaigners. While the concern expressed by the eighties campaigners must have had some general social impact, and certainly had an effect on me, the idea that there exists a widespread cultural objectification of women now seems to resemble one of those conceptual phantoms which, having once served as a basic structural element in our thinking (or in the thinking of some of us at any rate), vanishes completely when subjected to careful critical scrutiny.[4] Except that,

---

[4] As has arguably occurred with the traditional concept of mind, as a consequence of the influence of the views of thinkers such as Wittgenstein (Wittgenstein 1994), and Ryle (1949), for whom the traditional concept rests entirely on conceptual confusions.

in this case, the concept of objectification never seems to have been subjected to the sort of close scrutiny that might have been expected to generate such a result.

Since it would be a great pity if the concept of objectification were to be lost to ethical and political discourse on the basis of the unexamined assumption that it cannot be given a clear and useful sense, the time seems overdue for it to receive sustained philosophical attention. This is particularly the case for philosophers working in the area of applied and professional ethics, for this is an area in which one still regularly finds the concept invoked. One reads for example of the objectification of patients, particularly geriatric and terminally ill patients, by healthcare professionals.[5] One also comes across the idea that research subjects are at risk of objectification, particularly in biomedical research.[6] One hears less about, but surely has good reason to suspect, the presence of objectifying attitudes and behaviours in the criminal justice system, in connection with the various forms of institutional racism that investigators have detected there.[7] Finally, and closest to home for me, one surely ought to suspect the existence of objectifying attitudes and behaviours throughout the education system, where sexism and racism ought to be matters of permanent vigilance, and where the role and reliability of professional judgement becomes immensely important, particularly at advanced levels.

Accordingly, this book has its origins in my attempt to get clear on what objectification really is. However, this task turned out to be less straightforward than was initially imagined. It proved necessary, as my enquiry progressed, to ask why contemporary applied and professional ethics has not dealt very successfully with the concept of objectification hitherto. The attempt to answer this question led me first to consider an apparent crisis currently afflicting the principle of respect for persons in applied and professional ethics. This in turn led me to identify a distinct class of moral wrongs – which I term *interpretive* moral wrongs – that, to my knowledge, have not previously been explicitly distinguished, or theoretically analysed, as a class. Getting to grips with the only available theoretical material that seemed capable of shedding really useful light on the distinctive features of interpretive moral wrongs did, in the end, help with the initial task of understanding objectification. But it also led to something more – to what felt to me like a rediscovery of a dimension of our ethical obligations, the existence of which seems barely to be suspected in contemporary approaches to applied and professional ethics.

As a consequence of the unfolding character of the enquiry, this book embeds an interpretation of objectification in a broader narrative, which begins by considering the crisis currently afflicting the principle of respect for persons, and concludes by locating my analysis of objectification in terms of a more general ethic of *self-interpretation*.[8] The conclusion drawn is that the prospects for a revised and revived ethic of respect for persons are closely bound up with the prospects for providing a

---

[5]Code 1995.

[6]de Castro 1998.

[7]Bowling and Phillips 2001.

[8]This is the primary concern of Part II.

viable interpretation of objectification and related wrongs; and these prospects are in turn bound up with the rediscovery of the ethics of self-interpretation. I will argue that it is only by finding a way to bring the class of moral wrongs that essentially embody interpretations of their victims – of which objectification is a member – within the ambit of a contemporary ethic of self-interpretation, that we can rediscover the true contemporary relevance of the principle of respect for persons, and re-establish it on a secure footing.

1.2 The Structure of the Book

In order to prepare the ground for a more general enquiry into the ethics of self-interpretation, Chapter 2 investigates problems currently afflicting attempts to apply the ethical principle of respect for persons in contemporary applied and professional ethics. Although this principle has in the past been presented as *the* foundational ethical principle, and as being of particular relevance in applied and professional ethics, developments over recent decades suggest that a discrete principle of respect for persons may be increasingly irrelevant in practice. Recent work in biomedical ethics, animal welfare ethics, and environmental ethics tends to suggest that respect for persons is only of local significance – since the class of moral patients is by no means co-extensive with the class of persons. Where something analogous to a principle of respect for persons does survive, in the form of a principle of respect for autonomy, it has undergone a fundamental transformation. No longer a candidate for the foundational ethical principle, it is simply one element in the professional's ethical armoury, to be brought to bear piecemeal on the basis of what are in effect common sense ethical intuitions.

Chapter 3 considers an example of widely recognised ethical wrong that has the potential to prompt a critical revision of the picture of the continuing relevance of the principle of respect for persons that emerged from Chapter 2. The wrong in question is that of discrimination. I offer an analysis of discrimination that suggests that it cannot be reduced to any form of straightforward unfairness or injustice, since it includes an essential interpretive element. One can only be discriminated against *as* a member of this or that group. (By contrast, when one is bullied or exploited, one may be *simply* bullied or exploited.)

Chapter 4 discusses some potential counterexamples to the analysis of discrimination as an interpretive moral wrong. Consideration of these counterexamples leads me to identify stereotyping as an essential component in discrimination. Discrimination combines injustice with stereotyping, but stereotyping also turns out to be a significant interpretive moral wrong in its own right.

Chapter 5 focuses on the concept of objectification. Three 'stages' of objectification are distinguished. All three stages of objectification include an element of instrumentalisation. However, two of the three stages also involve an element of stereotyping, and thus qualify as interpretive moral wrongs. The latter forms of objectification accomplish a more comprehensive reduction of their victim to the

status of a mere means, since in these cases the victim is not only used, but is also encouraged to interpret herself as an object of use. When a victim of objectification interprets herself as an object of use the wrongs associated with straightforward instrumentalisation are perpetuated and intensified. The example of sexual objectification serves here to illustrate the nature and effects of objectification, thus understood.

Chapter 6 addresses some critical questions raised by the analysis of objectification in Chapter 5, and seeks to draw some more general conclusions from it. If objectification in its more developed forms involves an element of instrumentalising self-interpretation, and if (as in sexual objectification) the materials for such instrumentalising self-interpretations belong, inter alia, to certain common sense conceptions of who and what we are, a fully adequate analysis of objectification requires a critique of the relevant common sense beliefs and assumptions. In short, it must be radical, rather than merely reformist, in both its ambitions and its approach. I discuss and contrast reformist and radical approaches to the analysis of sexual objectification and related moral wrongs, including the Marxist concept of commodification, before going on to argue that the critique of common sense must also extend to include a critique of the common sense assumptions that apparently underlie certain scientific conceptions of human nature. This chapter, and with it Part I of the book, concludes with an examination of the normative basis of the analysis of interpretive moral wrongs. I argue that the instrumentalising or otherwise distorted self-conceptions on which interpretive moral wrongs depend not only serve to facilitate and perpetuate a range of offences against persons, but also undermine the dignity of their victims directly. Kant teaches us that human dignity is paradigmatically displayed in autonomous action. I argue that human dignity is also exemplified in undistorted self-interpretation, and thus in self-knowledge. A concern with undistorted self-interpretation should not replace a concern with autonomous action, but undistorted self-interpretation is arguably a presupposition of autonomous action. Notwithstanding the concerns raised in Chapter 2 then, the principle of respect for persons still has a significant role to play in an ethic of self-interpretation, whose distinctive concern is with interpretive moral wrongs.

In Chapter 7, which opens Part II of the book, I begin a search for sources for an ethic of respect for persons, as an ethic of self-interpretation, drawn from modern European philosophy. Hegel was perhaps the first philosopher to appreciate that an ethic of respect for persons makes sense only in the context of an ethic of *recognition*. I outline some salient features of the Hegelian theme of recognition, but also draw attention to its inherent essentialism – a feature that is unlikely to be regarded positively in contemporary applied and professional ethics.

Chapter 8 seeks to address the problem of the relationship between an ethic of self-interpretation and essentialism by turning to what we might refer to as the 'sceptical essentialist' tradition of post-Hegelian philosophy, which includes Heidegger, Sartre and Foucault. Heidegger's concept of authenticity embodies a rejection of essentialism, while simultaneously acknowledging that the critical questioning of who and what we are cannot cease to be a concern for us.

Although Heidegger sought to downplay the ethical relevance of authenticity, the concept does indicate a possible future for a sceptical essentialist ethic of self-interpretation. Nevertheless, it would suffer from two major drawbacks in such a role. Firstly, it looks to belong to an ethic of private perfection rather than one of public obligation. (In Kantian terms, it looks to be a source of imperfect duties, at best.) Secondly, while it has direct implications for the uncritical adoption of particular instrumentalised self-conceptions, it tells us little about the ways in which our actions potentially engender instrumentalised self-conceptions in others. These deficiencies are addressed in Chapter 9, through a consideration of Sartre's account of concrete relations with others. Sartre explores the way that struggles over instrumentalised self-conceptions pervade human personal relationships. He also shows how the adoption of an instrumentalised self-conception, or the inculcation of an instrumentalised self-conception in another, involves not simply a failure of authenticity, but an element of mendacity. It follows that we have a perfect duty not to adopt or inculcate instrumentalised self-conceptions.

Sartre's account of concrete relations with others points the way to a sceptical essentialist ethic of self-interpretation with real normative force. However, it remains excessively voluntaristic, and focused on the sphere of personal relationships. Chapter 10 aims to address these limitations by drawing on themes from the work of Michel Foucault. In some respects Foucault's thought marks a radical break with the phenomenological tradition to which both Heidegger and Sartre belong. But there are also clearly identifiable continuities between Foucault's concept of *subjection* and Sartre's concern with instrumentalising self-conceptions. In addition, Foucault focuses explicitly on the price paid for adopting a scientific self-conception – an issue of particular relevance for an ethic of self-interpretation geared to addressing issues in contemporary applied and professional ethics.

Chapter 11 returns to the key question of the normative foundations of a sceptical essentialist ethic of self-interpretation. The interim conclusion of Chapter 6 – that the normative foundations of an ethic of respect for persons as self-interpreting beings lie in the importance of an undistorted self-conception in a dignified human existence – might be regarded as vulnerable to the anti-essentialist views canvassed in Chapter 7. Specifically, the notion of an undistorted self-conception may seem to imply the existence of an essential human nature, to which we potentially have access. The sceptical essentialist tradition investigated in Part II might seem to evade such criticisms only at the price of losing all normative purchase. For example, Sartrean concerns about the mendacity involved in adopting or inculcating instrumentalised self-conceptions may seem groundless, if the self is the object of *invention*, rather than *discovery*. Chapter 11 argues that the characteristically Nietzschean virtue of honesty provides a slim but reliable basis for a sceptical essentialist ethic of self-interpretation. Rejecting the idea that our nature is simply (potentially) discoverable does not commit us to embracing the alternative view that all conceptions of human nature are simply inventions. It can still coherently be claimed that self-conceptions may be adopted honestly or dishonestly; and thus the notion of a distorted self-conception can be glossed as a self-conception that, if adopted, would be adopted dishonestly.

Chapter 11 concludes with a brief summary of the implications, for applied and professional ethics, of the expansion of the principle of respect for persons to include an ethic of honest self-interpretation. Professional roles typically involve the mastery of a body of authoritative knowledge. This knowledge may well become self-knowledge, in the person of the professional or his/her client. An adequate professional ethic of respect for persons would include an ethic of honest self-interpretation; and this would imply a duty to avoid inculcating (and, where practicable, to correct) dishonest self-conceptions, in both ourselves, and those whose lives we affect.

A note on style: There is an established convention in philosophical writing to use gender-neutral language, or feminine forms, wherever possible. I have adopted this as my default, but in some contexts (such as the discussion of master and slave in Chapter 7) it seems perverse to insist on using feminine forms. Accordingly, I have departed from the established convention where the context seemed to demand it.

References

Bowling, Ben, and Corretta Phillips. 2001. *Racism, crime and justice*. London: Longman.

Code, Lorraine. 1995. Persons, and others. In *Rhetorical spaces*, 83–102. London: Routledge.

de Castro, Leonardo. 1998. Ethical issues in human experimentation. In *A companion to bioethics*, eds. Helga Kuhse and Peter Singer, 379–389. Oxford: Blackwell.

Mackie, J.L. 1990. *Ethics: Inventing right and wrong.* London: Penguin.

Marx, Karl. 1967. Economic and political manuscripts of 1844. In *Writings of the young Marx on philosophy and society*, eds. Loyd D. Easton and Kurt H. Guddat, 283–337. New York, NY: Anchor Doubleday.

Ryle, Gilbert. 1949. *The concept of mind.* New York, NY: Harper and Row.

Sartre, Jean-Paul. 1989. *Being and nothingness* (trans: Barnes, H.). London: Routledge.

Wittgenstein, Ludwig. 1994. *Philosophical investigations.* Oxford: Blackwell.

Part I
Respect for Persons and Interpretive Moral Wrongs

Chapter 2
Fragmentation

Abstract The ethical principle of *respect for persons* has been considered by many to be central to any adequate system of ethics. Some have gone so far as to present it as the supreme principle of morality in general. However, this view looks increasingly untenable. Developments in applied and professional ethics over recent decades have challenged the centrality of respect for persons on a number of fronts. Firstly, the connection between ontological personhood and moral personhood has been widely and effectively questioned. Secondly, the connection between moral agency and moral patiency has been questioned. Finally, the principle of respect for persons has been displaced from its former centrality to occupy, in the form of a principle of respect for autonomy, a place in a range of mutually irreducible basic principles. Thus the fragmentation of the idea of the moral person has been accompanied by a decline in the apparent significance of the principle of respect for persons. In the light of this decline, it is reasonable to ask what future the principle of respect for persons can now have.

2.1 Respect for Persons, and Persons as Ends

The ethical principle of *respect for persons* has long been considered to be central to any adequate system of ethics. In their 1969 book *Respect for Persons*, R.S. Downie and Elizabeth Telfer go so far as to claim that the principle will serve as *the* supreme principle of morality in general.[1] However, in recent decades a series of difficulties have arisen in the areas of applied and professional ethics, which serve to place a significant question mark over the continuing ethical relevance of the principle, in its traditionally-understood form. In this book I seek to develop an interpretation of the ethics of respect for persons, which acknowledges the difficulties that currently beset the principle, but also recognises a distinct set of responsibilities toward persons that have been largely overlooked in recent debates. On this basis I will argue that the principle has a legitimate but restricted role to play, in a specific area of ethics. I will not be concerned, then, to defend the view of Downie and Telfer that the principle of respect for persons will serve as the supreme principle of morality. Instead I will be concerned with the continuing relevance of the principle in respect of a distinct class

---

[1] Downie and Telfer 1969, p. 15.

of moral wrongs (which I term *interpretive* moral wrongs), that have not previously received sufficient theoretical attention. Before embarking on any of this though, it is necessary to provide a sketch of what has traditionally been understood by an ethic of 'respect for persons'.

2.2 The Essence of 'Respect for Persons'

The *locus classicus* for an ethic of respect for persons is the ethics of Kant.[2] Kant considers human beings to be *ends in themselves* – that is, creatures with the capacity to pursue their own self-chosen goals, and to make rational choices in the pursuit of those goals.[3] Any creature that is an end in itself possesses a unique *dignity*, according to Kant.[4] And the inherent dignity of the end in itself entitles him or her to special forms of respect and consideration. Rationally self-determining ends in themselves are necessarily bound by firm canons of consistency and justice in their dealings with one another. They are obliged to treat each other only in ways that any such being could rationally consent to. In particular, they ought never to bring about, or permit, the *instrumentalisation* of any of their number – their reduction to the status of a mere means to another's ends. This requirement is best captured in the 'formula of humanity' variant of Kant's 'categorical imperative':

> So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, and never merely as a means.[5]

Respecting another as an end involves respecting her as a *self-determining* being. It thus involves respecting, within limits, her mature and considered choices, even – indeed especially – when we do not agree with those choices. It also involves ensuring that we do not frustrate or undermine the free development of this capacity for self-determination – again, even if allowing another freely to develop as a self-determining being promises to have uncomfortable and inconvenient consequences for ourselves, and relevant third parties.[6] A community whose members consistently acted in obedience to this principle would constitute what Kant terms a 'kingdom of ends'.[7] In the kingdom of ends each member treats every other as an end in herself, and never merely as a means or a resource.

---

[2]See Kant 1996.

[3]Ibid., p. 79.

[4]Ibid., pp. 84–5.

[5]Ibid., p. 80.

[6]Some of the relevant limits include: we are not required to regard the choices of others as paramount where they are likely to result in serious self-harm, or harm to others; nor are we required to regard them as paramount where they involve frustrating the choices of others; nor are we required to regard them as paramount where they are clearly based on ignorance or false beliefs.

[7]Kant 1996, pp. 83–4.

One way to bring out the distinctive character of the ethics of respect for persons thus understood is by comparison with another well-established and widely respected ethical principle, with which Kant's formula of humanity is clearly incompatible. Jeremy Bentham's principle of utility is probably the best-known candidate for a foundational ethical principle devised by a moral philosopher in recent centuries. As expressed by Bentham's fellow utilitarian John Stuart Mill, the principle runs as follows:

> [A]ctions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure and the absence of pain; by unhappiness pain, and the privation of pleasure.[8]

Bentham and Mill, the founders of modern utilitarianism, referred to this principle as the 'Greatest Happiness Principle'. Obedience to *this* principle requires us to act so as to maximise the happiness, or minimise the unhappiness, of all of those affected by our action. Adherents of a respect for persons ethic are however bound to hold that the principle of utility is not a reliable source of ethical guidance. It is difficult to altogether dismiss the idea that ethics must have some essential connection with the pursuit of happiness. [9] But considerations of happiness (or 'utility') alone are insufficient as a guide to our moral duty. Morality clearly requires us to act, on some occasions, in ways that promise to result in sub-optimal levels of happiness for those affected by our actions. For example, if someone is relying on us to be truthful, or to keep a promise, it seems *prima facie* wrong to deceive them, or break our promise, even where there is a strong chance that they will be happier in the long run if we do so. Regrettably perhaps, but apparently inevitably, the ethically permissible course, and the happiness-maximising course, sometimes part company.[10]

It is partly this insight (that we are sometimes required to act in ways that result in sub-optimal level of utility for those affected by our action) that underlies a commitment to the principle of respect for persons. The principle does not of course reduce to the denial of the principle of utility, and it certainly doesn't follow from the fact that obedience to the principle of respect for persons sometimes requires us

---

[8]Mill 1962, p. 257.

[9]This is not simply the view of utilitarians, such as Bentham and Mill, but is also fundamental to Aristotelian ethics (see Aristotle 1955, book one), and contemporary virtue theory. Neo-Aristotelians, such as Alasdair Macintyre, criticise Kant on the grounds that his theory divorces moral theorising from considerations of human flourishing (see Macintyre 1981, chapter 4).

[10]This is the point on which virtue theorists part company with utilitarians. In their view, that there is an essential connection between moral goodness and human happiness/flourishing does not entail that the maximisation of happiness can coherently be made the overriding moral goal. Kantians raise this to a point of principle: moral duty and happiness may sometimes coincide, but the pursuit of happiness cannot, in their view, be the goal of ethics. The four examples discussed in section 2 of Kant's *Groundwork for the Metaphysics of Morals* explicitly concern individual happiness and self-love. But each of them also bears on the question of duty versus the maximisation of general happiness, and in each case it becomes evident that, in Kant's view, moral duty and the pursuit of general happiness conflict (see Kant 1996, pp. 73–5).

to act in a manner that results in sub-optimal levels of utility for all those affected by our action that it must do so in all cases. Rather, what the principle expresses is the basic ethical conviction that we have duties to others as self-determining beings that are more fundamental than, and consequently override, any duties we may have to them simply as sentient beings, capable of experiencing pleasure and pain. In any case in which respecting someone's mature and considered choice will result in sub-optimal levels of happiness for those affected by the choice, but will not lead to any very significant harm, or to the frustration of the choices of others, the principle of respect for persons enjoins us to respect that choice.

Thus understood, the principle of respect for persons sits particularly well with the requirements of many professional codes of ethics, which demand that the professional maintain a relationship of loyalty and trust with the client. Whether in law, in healthcare, or even in education, the professional's duty is not well-captured by saying simply that she must act in line with the interests and/or preferences of all of those affected by her action. Rather, in her professional life, the lawyer has a specific duty to her client, which does not override her duty to the court, but does not extend immediately to the community as a whole either.[11] A doctor's duty to her patients may involve respecting the patient's choices even where they impact negatively on her health.[12] And a teacher's duty to her students may require her to respect and communicate truth, even in cases where the truth is unpalatable, and seems unlikely to contribute to the student's future happiness.[13] In all of these cases, in which professional duty and the pursuit of general utility diverge, the ethics of respect for persons provides us with a way to explain and justify actions that protect or foster self-determination, at the expense – potentially at least – of general social utility.

2.3 Contemporary Challenges

Nevertheless, all is not well with the ethics of respect for persons, even in the well-insulated world of professional ethics. It is not necessary to dissent from the core ideas informing the ethics of respect for persons as ends in themselves to entertain serious doubts about the applicability of the principle in contemporary contexts, particularly in institutional settings. The most important of the challenges confronting the ethics of respect for persons in practice are in my view well captured by Lorraine Code in her article 'Persons, and Others'.[14] Basing her discussion on a literary

---

[11]Carne 2010, pp. 63–5.

[12]See for example the British Medical Association 1993, sections 1:6–1:10.

[13]Thus we teach schoolchildren about the twentieth-century world wars not because we think it will make it them happier, or even (realistically) because we think it will make it less likely that similar atrocities will occur in the future, but because we think that understanding their history will help them better understand themselves and their world, and that such understanding is valuable in itself.

[14]See Code 1995.

account of the gradual disintegration of the personality of an elderly patient in a care home, Code considers how effective an ethic of respect for persons would be in protecting individuals in such circumstances, where their capacity for self-determination is in decline. The mentally defective, the mentally ill, those suffering from dementia etc. do not, Code notes, typically display the sort of consistent self-determining behaviour that is usually understood to be the hallmark of a 'person'.[15] They therefore represent cases in which, at least on a Kantian construal of what it is to be an end in oneself, the dividing line between 'ends in themselves' and mere 'things' has become unclear. Code does not suggest that there are circumstances in which we ought to consider dementia sufferers (for example) to be mere objects. Her point is that concentrating on protecting and fostering self-determination, as if that were the prime ethical concern in such cases, will not suffice to properly protect such individuals.

Code's discussion is interesting primarily for the richness of the critique it offers of the ethics of respect for persons. In order to fully appreciate this richness however, it is necessary to recognise that her concerns do not simply relate to the existence of a grey zone between the self-determining and the non-self-determining, into which (e.g.) care home residents may fall. They also relate to the range of features that go to entitle any being to moral consideration in the first place; and furthermore to certain distinctive features of persons that, in some circumstances, may leave them uniquely exposed to subtle forms of instrumentalisation.[16] Code's discussion implicitly identifies three major problems for a contemporary ethics of respect for persons, which for convenience I will label the problems of *integration*, of *moral standing*, and of *objectification*. In the remainder of this chapter I will explore each of these problems in turn. Taken together, I will argue, they represent nothing short of a crisis for a contemporary ethic of respect for persons.

2.3.1 The Problem of Integration

The problem of integration is connected with the various forms of supplementation that the basic principle of respect for self-determination requires in any context of personal or professional *care*. In her article 'Care and Respect' Robin S. Dillon argues that the Kantian conception of the forms of respect due to persons is excessively narrow, and that simply respecting the self-determining capacity of persons will tend to lead us to underestimate our obligations to them.[17] A strictly Kantian ethic of respect for persons may appear to function adequately between effectively autonomous individuals, but according to Dillon it fails as a general account of our

---

[15]Ibid., pp. 83–4.

[16]Ibid., p. 87.

[17]See Dillon 1992.

ethical obligations, because it fails to construe what it is to be an end in oneself sufficiently broadly.[18] Most evidently in contexts in which there is a professional duty of care, it is misleading to suggest that ethics primarily requires us to respect the self-determining choices of others. Rather, in such cases we have a clear and overriding duty to foster and protect their flourishing, regardless of how well-developed, or degraded, their capacity for self-determination may be.

Dillon's proposed solution to the perceived narrowness of the Kantian view of respect for persons is to appeal to something she calls 'care respect'.[19] To explicate this notion of care respect, Dillon turns to the conception of respect for persons developed by Downie and Telfer. As I have already implicitly suggested, Downie and Telfer's approach has many similarities with that of Kant. They argue that the principle of respect for persons will serve as the supreme regulative principle of morality, unifying and underpinning the ordinary rules and judgements of social morality.[20] By contrast with the Kantian approach however, Downie and Telfer build a teleological element into the very foundation of their concept of respect for persons. To respect a person as an 'end' is, on Downie and Telfer's view, to respect him or her for,

> [T]hose features which make him what he is as a person and which, when developed, constitute his flourishing.[21]

The initial attractiveness of Downie and Telfer's approach consists partly in the fact that by focusing on features that, when developed, constitute the flourishing of persons, rather than focusing more narrowly on the capacity for self-determination, they are able to incorporate elements of the Kantian view, while at the same time going beyond it, and including elements that it appears wrongly to exclude. Thus while Downie and Telfer insist, in Kantian fashion, that human persons are deserving of respect as the possessors of a rational will, they simultaneously emphasise the extent to which the activity of the rational will involves the exercise of capacities that are only tangentially connected with reason. If we were not sentient beings, possessed of desires and aversions, and also rule-following beings, able to determine ourselves to act in obedience to self-imposed principles, they observe, we would not be capable of purposive rational action at all. The various rational and non-rational capacities involved in rational willing go to make up what Downie and Telfer term 'personality'. And it is as the possessors of *personality*, paradigmatically as the initiators of complex practical projects, that human persons are held by them to be genuinely worthy of respect as ends in themselves.[22]

While Downie and Telfer are happy to align themselves with Kant then, in suggesting that human persons are worthy of respect as the possessors of a rational will,

---

[18]Ibid., pp. 72–7.

[19]Ibid., p. 73, and passim.

[20]Downie and Telfer 1969, pp. 15, 33, 38–64.

[21]Ibid., p. 15.

[22]Ibid., p. 29.

they also aim to go beyond Kant by arguing that the rational will cannot in practice be separated from personality, and respecting others as personalities involves acting so as to secure their flourishing. The most notable point of divergence between Downie and Telfer's position and that of Kant is, accordingly, the position they take on duties of beneficence. Kant distinguishes between the perfect and the imperfect duties of moral agents.[23] A perfect duty is any duty that we have an overriding obligation to act on, in all cases. Thus Kant holds that duties not to deceive, manipulate or instrumentalise others are perfect duties, obedience to which is an absolute requirement for all moral beings. Imperfect duties, by contrast, are duties the performance of which is meritorious, but not obligatory in all circumstances. Thus duties of beneficence and charity, which common sense ethical reasoning recognises as morally admirable, are imperfect duties, calling for moral praise, but by no means obligatory in all circumstances.[24] To respect another as an end is, for Kant, as we have seen, essentially a matter of acknowledging that others have the capacity to pursue ends of their own, and recognising that reason enjoins us to respect that capacity by not frustrating them in their pursuit of those ends. We are not, in Kant's view, obliged to *assist* others in the pursuit of their own ends in anything like the same way. Thus while the possessor of a moral will should recognise a fundamental inconsistency in any action that has the effect of systematically frustrating the projects of others, she will not necessarily find any such contradiction in failing to assist, in a particular case, the projects of another.[25] By contrast, Downie and Telfer argue that the principle of respect for persons does more than support imperfect duties of beneficence (along with perfect duties of non-interference and respect for self-determination). Respecting another human being as a person involves, they insist, making her ends our own.[26] It is not enough on this view simply to avoid

---

[23] Kant 1996, pp. 73, 512–22.

[24] Ibid., pp. 514–20.

[25] Strictly speaking, there is, for Kant, a contradiction in failing to assist others, but it is a contradiction of a less fundamental kind. Kant's distinction between perfect and imperfect duties is founded, as Allen W. Wood explains (Wood 1999, pp. 82–4), on a distinction among the ways in which the adoption of a subjective principle of action may turn out to be contradictory, and therefore rationally unacceptable. Some subjective principles are such that their adoption would give rise to what Onora Nell (O'Neill) (Nell 1975, chapter 1) terms a 'contradiction in conception'. Such a principle cannot be willed as a universal law because there could be no coherent system of nature structured by the relevant law. Other subjective principles are such that adopting them would give rise to a 'contradiction in volition' (Ibid). Here the principle could serve as the basis for a coherent system of nature, but the system in question would not be one with a place for human beings, as they are actually constituted. For example, there is a possible system of nature in which it is a natural law that nobody will help another when they need it – everyone will be ruggedly individualistic. This system would not be intrinsically self-defeating, but it is not a system that human beings as they are actually constituted could successfully inhabit. Human lives are marked by interdependence. Consequently, humans can conceive of, but cannot coherently *will*, the existence of such a system of nature. From this it follows, for Kant, that the duty to assist others in need is an imperfect rather than a perfect duty. To assist others in need is morally meritorious, but to fail to take an opportunity to do so is never straightforwardly wrong.

[26] Downie and Telfer 1969, pp. 25–30, 37.

using others as means; one is also obliged to actively sympathise with them, and to assist, where appropriate, in the realisation of their goals. Understood in this way, the duty to benefit others is on a par with our duty not to deceive or instrumentalise them, and the principle of respect for persons grounds what are in effect perfect duties of beneficence.[27]

As a consequence, the principle of respect for persons is, according to Downie and Telfer, able to play a unifying and foundational role with respect to both the ordinary rules and judgements of social morality, and the dominant contemporary traditions in normative ethics. The fundamental principles of both Kantian and utilitarian moral philosophies can, Downie and Telfer claim, be founded on the principle of respect for persons, properly understood. They go on to illustrate, with reference to the political values of liberty, equality and fraternity, how the principle of respect for persons can resolve apparent tensions between such positions, uniting classical utilitarianism, and Kantian respect for autonomy, under the umbrella of ethical and political liberalism.[28]

Downie and Telfer's project is evidently then an *integrative* one. Their aim is to show that, properly understood, the principle of respect for persons can serve as a basis on which to integrate apparently divergent traditions in contemporary philosophical ethics, and liberal political philosophy. This is what, for Dillon, makes their project so significant; and she observes that this integrative project is a distinctive task of contemporary feminist moral theorising.[29] But however keenly the need for such an integrative approach might be felt, the unfortunate truth is that the price of breaking away from the Kantian conception of respect for persons, and making the flourishing of others a paramount ethical goal in its own right, is to build very significant tensions into any associated ethic of respect for persons. There are powerful currents within contemporary applied and professional ethics that suggest that it is simply not possible to supplement a principle of respect for self-determination with a principle of 'care', in an integrated ethic of respect for persons.

One effective way to bring out this problem of integration is to focus on the phenomenon of paternalism, which is a major concern in contemporary biomedical ethics (though it is certainly not restricted to medical ethics contexts).[30] Paternalism is the well-known syndrome of 'the expert knows best', which threatens to break out wherever the expert knowledge of the medical professional collides with the beliefs and values of the patient. Paternalism is of concern partly because experts are not infallible. The expert may think she knows best in a given case, but it is conceivable that she may be mistaken – in a given case it *could* be that the patient's judgement is in fact the more reliable guide (either because the patient has superior insight into the details of the case, or purely from luck). However, concern about paternalism does not stem primarily from concerns about the fallibility of experts. It may be

---

[27]Ibid., pp. 28–9.

[28]Ibid., chapter 2.

[29]Dillon 1992, p. 73.

[30]Beauchamp and Childress 2001, pp. 176–94.

that, in a given case, the expert is wrong; but she is usually much more likely to be right (she is, after all, the expert!). The real core of the concern with paternalism lies in the fact that, even where the expert does know best, the competent patient still has a right to be involved in the decision-making process, to the extent of overruling the expert in certain circumstances. For example, a doctor may advise on a treatment regime, but if a competent patient disagrees over what constitutes a tolerable level of discomfort, or has (say) ethical or religious qualms about the nature of the treatment on offer, the final decision may well have to be hers.

The phenomenon of paternalism is instructive in the present context because in situations in which worries about paternalism are pertinent we are required to take a position on whether, in the case before us, respecting individual self-determination is more important than securing individual well-being. As we have already seen, utilitarians hold that the maximisation of collective well-being is the fundamental concern of ethics. Provided the patient's potential unhappiness with a particular decision can be factored into the utilitarian calculus then, it is theoretically possible to be consistently utilitarian about any such choice: it may be that the patient doesn't like what the doctor is recommending, but her present unhappiness may easily be outweighed by her long term satisfaction with the superior health benefits that will ultimately result. It is hard to believe though that nothing of ethical importance is left out when the patient's unhappiness with the decision is accounted for in this way. Intuitively, it seems vital to find some way to acknowledge that the ethical challenge that arises when the patient's own choice stands to be overruled cannot be satisfactorily dealt with by factoring how unhappy she feels about being overruled into our calculations.

This intuition can be backed up by considering the implications of withholding key information from the patient. From the utilitarian standpoint, if the patient's unhappiness with the clinically-indicated course of treatment can be ameliorated, or even avoided altogether, by withholding key information, then it may well be admissible (even, strictly, obligatory) to do so. But this seems to completely misinterpret the nature of the moral conflict in any such case. If it is right to involve the patient in the decision-making process at all, then it is surely important to ensure that she is equipped with the information she needs to participate properly. The consequences for her happiness of ensuring that she is properly informed are then beside the main point; and an ethical approach that fails to register this fact is to that extent defective. If the patient has a right to be involved in decisions about her own care, this right must be substantially independent of considerations pertaining to her happiness.[31] It follows that conflicts in such situations cannot be dissolved by

---

[31] Plato's apparently rather light-hearted example of the doctor who is used to treating slaves, who pours scorn on the 'gentleman' doctor, who seeks to inform his patients about their condition, at the same time as providing a cure, provides an illuminating alternative perspective here. While the main medical goal may be to achieve a cure, it is readily apparent that: (1) some 'cures' necessarily involve education (e.g. concerning diet and exercise), and (2) what might once have been considered acceptable in the treatment of slaves is by no means appropriate for the treatment of free citizens (Plato 1970, pp. 362–3.).

factoring the patient's unhappiness into a utilitarian calculus. That such situations do strike us as embodying genuine moral conflicts suggests that it is ethically indispensable to have some recourse to a principle of respect for self-determination.

That said, such situations simultaneously serve to highlight the difficulties that would arise if the principle of respect for self-determination were itself to be installed as *the* supreme principle of ethics. If it were simply a matter of supplying the patient with the relevant information, and allowing her to make up her own mind, the sense of a moral conflict in such situations would again disappear. But it is precisely because we feel the patient has a right to be properly involved in such decisions, *and* that it may sometimes be right for the doctor to overrule her, in her own best interests, that we feel that such situations embody genuine moral conflicts. Just as it would be wrong to approach such situations as if the pursuit of individual or collective happiness must always override other considerations, it would be equally wrong to approach them as if the principle of respect for self-determination should always have ultimate authority.

The problem of integration is not reducible to the problem of paternalism, but the problem of paternalism highlights vividly why integration is a problem. Concerns about paternalism only arise because the pursuit of well-being and the protection of self-determination sometimes pull us in different directions at the same time. And yet it seems that we cannot ignore either consideration. Thus the principle of respect for self-determination needs somehow to be complemented, and counter-balanced, by a principle that enjoins the pursuit of individual well-being, and the avoidance of harm.

It was partly in order to address such problems that Tom Beauchamp and James Childress developed the 'four principles' approach to biomedical ethics, some decades ago.[32] Beauchamp and Childress argue that any adequate approach to biomedical ethics must be based on more than one fundamental principle. Indeed, according to Beauchamp and Childress four fundamental principles are required: a principle of *respect for autonomy* (that is, respect for the autonomous choices of patients), a principle of *non-maleficence* (not to cause of harm), a principle of *beneficence* (doing good where possible), and a principle of *justice* (fairness in apportioning harms and benefits).[33] The four principles approach appears to be the outcome of an attempt to thresh the key insights of the main traditions of modern moral theory from the more problematic elements that surround them, and to articulate these insights in a form that, while addressed specifically to professionals, simultaneously projects a clear picture of patients' rights. It aims to occupy a mid-level position between apparently free-floating professional ethical codes and philosophical moral theories, giving medical professionals some insight into why each of the four principles is so important, while avoiding enmeshing them in the deeper problems that afflict more theoretically ambitious approaches to ethics.[34]

---

[32] See Beauchamp and Childress 2001.

[33] Ibid., chapters 3–6.

[34] Ibid., chapters 1, 8.

At first sight the differences between the four principles approach and that of
Downie and Telfer may appear slight. The former may appear to be essentially
a codification of the basic requirements of the ethics of respect for persons, as
presented by Downie and Telfer. But closer inspection reveals a very real differ-
ence between the two approaches, which ultimately opens into a chasm. It is a
key element in Beauchamp and Childress's view that the four principles must be
'balanced' against each other.[35] Thus in a case in which paternalism seems a threat
we must balance the requirements of the principle of beneficence (the pursuit of
well-being) with those of the principle of respect for autonomy (respect for individ-
ual self-determination).[36] It is initially tempting to take this metaphor of balancing
as implying that we should think in terms of 'weighing' the two principles against
each other. But further investigation reveals that, on Beauchamp and Childress's
model, balancing of *this* sort is quite impossible. The four principles are held to be
independent and irreducible. Thus they will not admit of accurate direct comparison
with each other, or accurate indirect comparison with a single independent theoreti-
cal measure. Consequently, no such process of weighing is possible.[37] This doesn't
necessarily make Beauchamp and Childress's talk of 'balancing' misleading. The
notion of balancing they have in mind may be better thought of on the model of
the checks and balances built into US constitutional arrangements – each princi-
ple is supposed to be able to override the others in certain circumstances, while
being overridden by them in other circumstances, much as Congress, the Supreme
Court, and the Executive may all overrule each other, depending on the circum-
stances. Thus one may 'balance' the four principles not by weighing them on some
notional scale, but by a process of limit-setting: each principle potentially limits
every other.[38] But it then follows that the basis of this balancing must be essen-
tially intuitive – there can be no higher order principles for applying the principles.
(Though balancing will require careful specification of the principle in a given case,
and thorough knowledge of the relevant facts of the case.)

With their four principles approach Beauchamp and Childress claim to be reflect-
ing principles inherent in the 'common morality', which all 'morally serious'
persons share.[39] This may sound rather similar to Downie and Telfer's ambition
to present a unified picture of ethics that will support the ordinary rules and judge-
ments of social morality. But the ambitions of the four principles approach are in

---

[35]Ibid., pp. 18–21.

[36]Ibid., pp. 176–94.

[37]At the theoretical level, Beauchamp and Childress concede, commensurability is impossible:
'we reject. . . the hypothesis that all leading principles in the different major moral theories can
be assimilated into a coherent whole'; 'In this "theory" there is no single unifying principle or
concept' (Beauchamp and Childress 2001, pp. 338, 405).

[38]I should point out that Beauchamp and Childress's initial presentation of the process of balancing
(Ibid., pp. 18–21) very much suggests that they do have the weighing model in mind. The inter-
pretation of balancing hazarded here is motivated by the recognition that other elements of their
position seem explicitly to rule out such weighing.

[39]Beauchamp and Childress 2001, pp. 2–5.

fact much more modest. For Downie and Telfer it is important that ordinary rules and judgements of social morality, such as moral rules of thumb having to do with beneficence and respect for autonomy, should be shown to be grounded in the principle of respect for persons, and in a particular conception of human flourishing. Their teleological approach to respect for persons aims quite self-consciously to unite deontological with teleological ethical themes. The view stands or falls with the tenability of the claim that the supreme principle of respect for persons can successfully embody both respect for the self-determination of persons and promotion of their well-being. And this is just what Beauchamp and Childress implicitly reject: for them there is no ruling principle, supreme over the four principles of respect for autonomy, beneficence, non-maleficence, and justice. One cannot appeal to an independent principle that will indicate when, for example, we are to prefer respecting patients' choices to promoting their well-being. And this is no arbitrary departure on their part. It is apparently forced on them by certain key ethical phenomena with which their model must deal – such as the problem of paternalism. For Downie and Telfer, potential conflicts associated with concerns about paternalism are always in principle resolvable by appeal to the supreme principle of respect for persons. But according to Beauchamp and Childress there can be no neat theoretical resolution of such dilemmas. Such dilemmas *are* resolvable in practice, in their view, but *not* by appeal to some supreme theoretical principle, with the aid of which the competing claims of beneficence and respect for autonomy might be rendered commensurable.

The conclusion we must draw is that the project of presenting an integrative contemporary ethic of respect for persons has encountered serious difficulties. The integrative ideal remains important within, for example, feminist moral philosophy, as well as more generally. But the remarkable success and widespread influence of the four principles approach serves to illustrate just how difficult the task of integration is likely to be. There is, it seems, no coherent and unified set of features of persons that we are in all cases required to respect. Rather, respecting the various different features of persons that call for respect promises to potentially pull us in different directions at once. Or, if it is insisted that the features of persons that (as Downie and Telfer put it) constitute their flourishing will form (when developed) an integrated whole, then it seems we may as well say that the person-as-a-whole has turned out to be something that the ethics of respect for persons looks currently unable to deal with. On Beauchamp and Childress's view it is necessary to abstract certain features of persons that command respect. And in abstracting and seeking to respect these different features we find we are often subject to contradictory demands. For the purposes of moral theorising, the person-as-end has become notably fragmented.

2.3.2 The Problem of Personhood

In addition to the apparent fragmentation that the idea of the moral-person-as-integrated-whole has undergone in recent decades, a second, related, fragmentation has occurred, affecting the contemporary conception of *what it is to be* a person. In

a rather curious way, this second fragmentation seems to be a consequence of certain tendencies inherent within the principle of respect for persons itself, as hitherto interpreted. This is evident in the way that the fragmentation in question begins with the attempt to actually apply the principle, in areas of urgent contemporary ethical concern.

One significant ethical application of the idea that the principle of respect for persons will serve as the ultimate ground for the ordinary rules and judgements of social morality is in bioethics, in relation to particularly recalcitrant difficulties around beginning- and end-of-life decisions (abortion, euthanasia etc.). Standardly, the deliberate taking of human life is regarded as a prima facie moral wrong. This attitude finds expression in cherished ideas, shared by people from many different cultural and moral traditions, concerning the sanctity of human life. However in the case of abortion, and also, in certain circumstances, in cases of euthanasia, many people feel that the deliberate taking of human life may sometimes be straightforwardly permissible (that is, as something other than the lesser of two wrongs). Finding a way forward in the face of these questions requires us to step back from our immediate emotional reactions, and search for some principle that will help us navigate the maze.

A possible way to do this would involve de-coupling our duties to others from the fact of their biological humanness, and attaching them instead to human 'personhood'. If this could be done, then in cases in which we are faced with the possibility of ending a human life, but where personhood is either entirely undeveloped, permanently absent, or irrevocably lost, we would seem to have some basis for thinking that ending that life is straightforwardly permissible. Michael Tooley is a well-known advocate of such an approach.[40] Tooley acknowledges that what personhood consists in is a controversial matter, and there is thus an inevitable looseness attaching to the concept.[41] Nevertheless, he argues, it is capable of doing important bioethical work. We tend to associate personhood with the higher mental functions of human beings; and, observing that the deliberate and wanton destruction of all of the higher mental functions of a human individual human being would elicit much the same moral disapprobation as would the murder of that individual, Tooley concludes that it is not human life that properly commands ethical respect, but personhood.[42] It would follow that where personhood is absent, the deliberate termination of biologically human life is in some circumstances permissible.

For Tooley's strategy to succeed however, it will be necessary to spell out in some detail quite why respect for personhood, as opposed to mere biological humanness, matters. This requires us to focus closely on the criteria for personhood. And this is a high-risk strategy. The possibility of separating personhood from biological humanness is not one that Downie and Telfer, for example, seem particularly eager to explore. For at least some of the time they seem to prefer to treat 'personality'

---

[40]See Tooley 1998.

[41]Ibid., pp. 120–1.

[42]Ibid., pp. 117–9.

as primitive. But the use that Tooley wishes to make of the idea of personhood as the locus of respect requires a different strategy, since to treat personhood as primitive would be to remain entangled in the existing ambiguity that attaches the label 'person' indifferently to biological humans and (e.g.) rational agents.[43] As soon as we cease to treat personhood as primitive though, and begin to set out a criterial definition of personhood, the concept begins to unravel.

Consider for instance two key features of persons, discussed by Downie and Telfer, and by Tooley: sentience, and a capacity for rational deliberation and principled decision-making.[44] It seems prima facie reasonable to think that sentience fits us for certain forms of moral consideration. But why should sentience be essentially linked to other facets of personhood (traditionally conceived), such as a sense of our own identity? In his second *Meditation* Descartes systematically doubts his own sentience, and in doing so he takes himself to have moved toward a clearer sense of his own identity.[45] On this highly influential view, sentience clouds and potentially disrupts our sense of our own identity. And even on a much more moderate view of the relationship between sentience and personal identity we could reasonably ask: why should our being sentient contribute anything one way or another to the maintenance of a sense of our own identity? A capacity for rational deliberation and principled decision-making on the other hand, where it is thought of as involving not simply conformity to rules (in the sense of 'conformity' in which natural phenomena conform to the laws of nature), but *obedience* to rules, does seem to have direct implications for personal identity. To obey a rule one must understand oneself as an addressee of the rule, and for this to be the case one evidently needs a sense of personal identity.

However, while the capacity for rule-following is evidently closely linked with personal identity, it doesn't seem particularly closely linked with whatever it is that fits us for moral consideration. It is clear enough why a capacity for rule-following might be thought to be linked to moral *responsibility* – if moral action is essentially about principled decision-making, then it is a necessary prerequisite of having moral duties that one is capable of obeying moral rules. But why should a capacity for rule-following be thought to be particularly closely connected with our entitlement to moral consideration? (Otherwise put: why should our moral responsibilities extend only to those capable of obeying moral rules?)

The above two examples of commonly accepted features of personhood which, nevertheless, appear to have widely differing implications for personal identity and moral considerability, suggest that personhood is far from being the monolithic phenomenon Tooley's discussion assumes. As soon as we begin to specify and examine the various favoured criteria for personhood we find that we are pushed to acknowledge a further ambiguity in the concept of a person. The concept is not simply, as Tooley notes, ambiguous between members of the species *homo sapiens*, and

---

[43] Ibid., p. 117.

[44] Downie and Telfer 1969, pp. 21–2; Tooley 1998, p. 120.

[45] Descartes 1986, pp. 19–20.

'individuals who enjoy... the type of mental life that characterises normal adult human beings'.[46] Rather, if the set of criteria for personhood, insofar as personhood relates to moral considerability, are distinct from the set of criteria for personal identity, we need to further distinguish between the *ontological* personhood of beings, of whatever species, and their *moral* personhood.[47] Ontological personhood is a matter of possessing certain qualities traditionally associated with distinctively human consciousness and agency. It is connected with the possession of some more-than-rudimentary sense of self, with all that that entails for the subjective dimension of one's existence, with respect to our possession of preferences and interests, for deliberate action, for having a subjective personal history, and so on. The ontological person possesses a developed form of self-awareness – she is the subject-object of self-understanding and self-knowledge. Ontological personhood has usually been held to rest upon some form of psychic continuity. (What is it, for example, about the identity of persons that makes the question whether the subject of a biography is the same person at the beginning and the end of the biography seem gratuitous, in a way that the question whether the nation to which she belonged was the same does not?) With respect to the metaphysics of personhood, while Hume's observation that the self is never to be found among the succession of psychic appearances has its attractions, philosophers have on balance tended to be more impressed by the Kantian response that it is a mistake to go looking for the self among the appearances at all, since *selfhood* is precisely that in virtue of which we experience a succession of psychic appearances as being 'ours'.[48] These metaphysical considerations aside, ontological personhood is also thought of as that which is progressively lost in degenerative mental conditions such as Alzheimer's.[49] Thus for all the philosophical complexities surrounding the notion of personal identity, the fact that we can empirically distinguish between individuals whose identity has suffered some form of severe and more or less permanent deterioration, and others whose personal identity has remained intact over time, suggests that we can use the term with some confidence.

Moral personhood, by contrast, concerns not so much what one is, as what one qualifies as – it is a matter of one's entitlements. Etymologically, the term 'person' derives from the Latin *persona*, which originally meant a mask used in a theatrical performance. In keeping with this origin, personhood in the moral sense is essentially a matter of being a moral 'player'.[50] Moral personhood, analogously with legal personhood, is not simply a matter of what one is, but of one's acknowledged or merited *status*. Even assuming we had the answers to all the puzzling metaphysical questions concerning personal identity then, we would not necessarily

---

[46]Tooley 1998, p. 117.

[47]See Beauchamp 1999.

[48]Hume 1969, pp. 299–303; Kant 1929, pp. 151–5.

[49]Tooley 1998, p. 122.

[50]The *locus classicus* for the use of the term 'person' in the moral and political sense is Locke 1924.

have an answer to the question: what is it about ontological persons that fits them for moral consideration? One may certainly talk of individuals 'qualifying as' both ontological and moral persons, meaning by this that they meet both relevant sets of criteria. But the manner in which one qualifies as a moral person is evidently different from that in which one qualifies as an ontological person.

Once the formerly foundational concept of personhood is subject to such scrutiny, it is apparent that there is nothing natural or obvious about the idea that ontological personhood should serve as our criterion of moral personhood. It certainly could play such a role, but that it should do so is a point that stands in need of independent argumentative support, and which cannot be allowed to rest simply on the observation that many people are content to have it play such a role. This is bad news for views, such as that of Tooley, which appear to infer moral personhood from our intuitions concerning the destruction of ontological personhood without much in the way of intervening argument.[51] From the fact that most people think it important to provide ethical protection for ontological persons *qua* ontological persons, rather than *qua* biological humans, it does not follow that we ought only to provide such protection for ontological persons (or even, strictly, that we ought to provide such protection for ontological persons at all).

Contemporary ethicists cannot therefore afford to assume that ontological personhood is unproblematic as a criterion for moral personhood. And this is in part because the concept of moral personhood has *itself* become fragmented. The fragmentation of personhood, which occurs when concepts of personhood or personality, formerly treated as primitive, are subjected to close scrutiny (for instance as part of the process of questioning the common sense and supposedly 'intuitive' reactions people have in respect of controversial moral questions in contemporary bioethics) does not stop with the cleavage between moral and ontological personhood. Moral personhood itself has been shown to fragment under pressure. As we have seen, the separation of moral and ontological personhood creates difficulties for those wishing to argue that our primary ethical obligations extend exclusively to ontological persons. At the same time however, this separation is more or less unavoidable once we de-couple personhood from biological humanness, since the very act of de-coupling belongs to a more critical approach to what makes persons, and in the ordinary run of things biological humans, both the possessors of a personal identity and the proper objects of moral consideration. But a further consequence of the de-coupling of personhood from biological humanness is that, in addition to suggesting that some humans might not in fact qualify as persons, it raises the complementary possibility that some persons might not be humans (that is: some non-humans may qualify as persons). The possibility that some persons might not be humans is one that Tooley acknowledges, but doesn't pay a great deal of attention to.[52] For some, the idea may seem to conjure up science fiction possibilities – interesting at the level of theory perhaps, but of little practical significance. However for philosophers working in the area of animal welfare ethics, the

---

[51]Tooley 1998, p. 118.
[52]Ibid., 1998, p. 125.

possibility that some persons may not be humans, and that we may consequently have direct moral obligations to non-human animals, is a matter of urgent concern.

Depending on how moral personhood is construed, both advocates and opponents of the claim that we have direct moral obligations to non-human animals might be suspicious of the air of anthropomorphism that initially surrounds the idea that some non-human animals may qualify as persons. Opponents will no doubt suspect that our perception of the non-human animals in question is likely to be distorted by the comparison, and that, as a consequence, moral sympathies for animals are likely to be unreasonably extended. Proponents on the other hand will tend to be concerned that animals should be recognised as being worthy of moral consideration in their own right, rather than on the basis of their possession of some set of human-like qualities. However, provided we adopt a suitably critical attitude to the concept of personhood itself, anthropomorphism need not be a danger. Once claims about moral personhood are properly distinguished from claims about ontological personhood (some of which may of course turn out to be prerequisites for moral personhood), there need in principle be no anthropomorphism involved in attributing moral personhood to non-human animals.

Among those who have pressed for a bioethical framework that rejects the traditional view of the sanctity of human life, and of biological humanness as the criterion for moral consideration, none has been more influential than Peter Singer.[53] Singer has been prominent both in bioethics, narrowly conceived, and in animal welfare and environmental ethics. The main focus of Singer's efforts in animal welfare ethics has been the struggle against speciesism, understood as the arbitrary preference for members of one's own species in ethical matters.[54] Singer's attack on the notion of the sanctity of human life is of a piece with his campaign against speciesism, since the idea that all and only human lives possess a distinctive sanctity looks to be a good example – indeed the central example – of such an arbitrary preference.

Singer argues that treating non-human animals as worthy of moral consideration requires nothing more than an application of utilitarian orthodoxy.[55] The key question when determining the moral considerability of an entity is not whether it is capable of complex mental activities, but whether it is capable of suffering. And Singer is able to cite no less an authority than Bentham himself in unequivocal support of this view:

> What else is there that should trace the insuperable line? Is it the faculty of reason, or perhaps the faculty of discourse? But a fullgrown horse or dog is beyond comparison a more rational, as well as a more conversable animal, than an infant of a day, a week, or even a month, old. And suppose it were otherwise, what would it avail? The question is not, Can they *reason*? nor Can they *talk*? but, *Can they suffer*? [56]

[53]See Singer 1995; 1993; 2002.

[54]Singer 1995, pp. 1–9.

[55]Singer 1995, chapter 1; 1993, chapter 3.

[56]Bentham, quoted in Singer 1993, p. 57.

   Thus on Singer's utilitarian view, every being with a capacity to suffer is properly a member of the moral community, and every such being –human or not – is entitled to have its suffering taken equally into account.

   From the fact that we have a duty to take the suffering of all equally into account, it does not follow, however, that all suffering must weigh equally in the utilitarian calculus. Suffering evidently comes in degrees, and different species have different sensitivities.[57] What matters in Singer's view is not the nature of the painful event, but the nature of the ensuing suffering, and this will not be felt by all species equally. This does not mean though that animal suffering must always be downgraded compared to that of humans, even where the biological constitution of the animal involved makes it relatively insensitive to physical pain. Sometimes superior intellectual capacities can serve to reduce our suffering, compared with that of a member of another species in a similar situation. For example, a wild animal caught in a trap may suffer more than a human in a similar situation just because it cannot be given to understand that it will shortly be set free.[58] On the other hand of course, understanding the reason for one's discomfort can be a cause of intensified suffering, as when one realises that a mildly irritating physical symptom indicates the onset of a serious illness. Since suffering comes in many different forms, what matters, according to Singer, is that we take all suffering properly into account, within the limits of our imperfect knowledge. In particular, we should not allow the mere fact of being human to load the scales one way or another. Taking the suffering of each and every sentient being into account means assessing it with reference to its specific intensity, allowing for both the physical and the mental sensitivities of the being experiencing it.

   From the fact that Singer, like Tooley, opposes the traditional doctrine of the sanctity of human life then, it does not follow that he favours replacing biological humanness with ontological personhood as the criterion of moral consideration. For Singer it is the basic and biologically widespread property of sentience that is ethically most significant, and not the relatively exalted capacities usually associated with ontological personhood. Singer does not advocate a re-assessment of personhood, in either its moral or ontological senses. His approach is more an eliminativist one, which aims to suggest that we would do better not to speak of personhood in connection with ethics at all, since the term is always likely to carry speciesist overtones. However, provided we are sufficiently rigorous in distinguishing conceptually between moral personhood, ontological personhood, and biological humanness, this does not seem to be an insurmountable danger. We may then, with caution, continue to use the term; and it seems we must conclude that Singer's proposed extension of

---

[57] Singer gives the example of a slap that would be very painful for a human child but would be much less painful for a horse (Singer 1993, p. 59).

[58] Ibid., p. 60. The discomfort of a vaccination injection would be another case in point: the human adult, who understands why the injection is being given, is likely to suffer less, all things considered, than the animal – or even the young child – who doesn't understand the reason for this painful intervention.

moral considerability to all sentient beings requires (though he would not put it this way) a reform of traditional notions of moral personhood.

In addition to implying that moral personhood needs to be de-coupled from ontological personhood, Singer's views also have the effect of highlighting the fact that moral personhood is a more complex notion than might initially have been thought. Moral personhood also seems to combine two conceptually distinct spheres, namely the sphere of moral agency, and the sphere of moral patiency. Those falling within the former sphere are the subjects of ethical duties. Those falling within the latter sphere are the objects of those duties. But these spheres need not necessarily coincide in practice. All sentient humans and non-humans will qualify as moral patients, on Singer's view. But only humans will (so far as we know) qualify as moral agents. Most normal humans will occupy the intersection of the two spheres. But it is perfectly possible that under certain circumstances we may fall outside of the sphere of moral agency and possibly also outside of the sphere of moral patiency. While we may as yet have no reason to doubt that all moral agents are also moral patients, Singer's arguments suggest that we have good reason to think that many moral patients are not also moral agents.

Thus far we have followed the fragmentation of personhood through the separation of moral personhood from ontological personhood, and the separation of moral agency from moral patiency. The process of fragmentation does not end even here though, since it also seems possible to identify more than one 'tier' of moral patiency. At much the same time as Singer was developing the utilitarian case for animal welfare ethics, Tom Regan was developing a rights-based approach in the same area.[59] Regan's approach suggests that sentience and sophisticated self-awareness may not be the only factors we need to consider when attempting to determine the extent and nature of moral patiency. Rather than focusing on sentience as a criterion of moral patiency, Regan notes that, while the capacity for sophisticated self-awareness may not extend further than the human world, many non-human animals nonetheless possess some limited selfhood, understood as the capacity to make deliberate choices and pursue projects of their own. According to Regan, most mammals and birds are the experiencing 'subjects-of-a-life'.[60] As such they deserve more protection than would be afforded by Singer's utilitarian animal welfare ethic. Narrow human self-interest apart, there is no good reason, on Regan's view, not to show the same sort of respect for the choices of those non-humans capable of making meaningful choices as we do for those of human beings. If it is wrong to systematically thwart the choices of human subjects-of-a-life, by subjecting them to purely instrumental treatment, then it is wrong to systematically thwart the choices of non-human subjects-of-a-life. Human and non-human subjects-of-a-life are therefore entitled to protection both from unnecessary suffering and from instrumentalisation and exploitation.[61]

---

[59] See Regan 1984; 2003.

[60] Regan 2003, p. 93.

[61] Regan 1985, pp. 24–5; 2003, pp. 95–6.

It should be emphasised that although both Regan and Singer favour the extension of moral consideration to the non-human world, their views are by no means wholly compatible. If adopted, Regan's recommendation that we refuse to instrumentalise non-human subjects of a life just as we refuse to instrumentalise human subjects of a life would mean an end to the straightforwardly exploitative treatment of many non-human animals, and would involve, amongst other things, an absolute ban on many forms of factory farming, and animal-based product testing and research. Singer on the other hand would be prepared to tolerate such treatment in cases in which it appeared to be justified on the basis of the relevant utilitarian calculations. Despite their differences, the two types of view jointly suggest a further fragmentation: moral patiency may be of a kind that rests on sentience alone, or may be of a kind that rests on being the subject-of-a-life, which is not necessarily found connected with moral agency, but which does place a distinct set of demands on moral agents.[62]

Further reasons to think that Regan's category of subjects-of-a-life who deserve protection from exploitation and instrumentalisation, rather than simply consideration for their pleasures and pains, may mark out a distinct sphere of ethical concern, are suggested by certain contemporary developments in biomedical ethics. Onora O'Neill has discussed the importance of the distinction between mere self-determination and 'autonomy' in the Kantian sense, in connection with Beauchamp and Childress's principle of respect for autonomy.[63] O'Neill argues that biomedical ethics is in practice oriented more toward the former type of autonomy, which she refers to as 'individual' autonomy, to the exclusion of the traditional Kantian alternative – 'principled' autonomy.[64] She expresses her concern that this focus on respect for individual autonomy has a corrosive influence on the professional-client relationship, and tends to undermine the key element of trust, by suggesting an oppositional model of the basis of the relationship – even capricious and arbitrary choices must be respected, provided they are suitably informed.[65] It seems to be an open question whether, on O'Neill's account, all humans must be considered capable of exercising principled autonomy, or whether some of them (children, the mentally disadvantaged etc.) might be capable of individual autonomy only. But if the two forms of autonomy really are distinct, and if individual autonomy represents the limit of what is attainable for some humans, O'Neill's distinction lends further support to the idea of a fragmentation of the sphere of moral patiency. Some moral patients are such in virtue of their capacity for principled autonomy, but others may be moral patients in virtue of their capacity for individual autonomy only (though of course other factors such as sentience will also be involved). This is broadly consistent with the picture suggested by Regan, according to which we have duties to all

---

[62]Regan emphasises that to be a subject-of-a-life is to possess equal moral worth to other subjects-of-a-life, but not necessarily to be a moral person in the full sense (Regan 2003, pp. 93–4, 102).

[63]See O'Neill 2002.

[64]Ibid., chapter 4.

[65]O'Neill 2002, sections 4.1, 5.2.

beings capable of making meaningful choices which are independent of any duties we may have in virtue of their mere sentience, or indeed any we may have in virtue of their capacity for the more sophisticated forms of autonomy that may be achieved by some, though not all, moral persons.

Neither Singer nor Regan advocates the extension of orthodox moral personhood to non-humans. But ultimately it seems unimportant whether we choose to describe Singer and Regan as recommending the extension of forms of moral personhood beyond the human species, or whether we describe them as recommending that moral consideration be extended to non-persons. The key point for our purposes is that the line of thought that they and their colleagues in the fields of animal welfare and environmental ethics explore turns out to further fragment the traditional view of an ethic of respect for persons. Any concept of personhood founded on something analogous to Downie and Telfer's concept of *personality* will be much too insensitive to do justice to the range of concerns that arise following the de-coupling of personhood from biological humanness. And the process of refining and modifying the concept to tailor it more effectively to the various contexts in which it must be used serves to cast doubt on the continuing usefulness of the concept altogether. By the time we have distinguished ontological personhood from moral personhood, moral agency from moral patiency, principled autonomy from individual autonomy, and the moral considerability of choice-makers from the moral considerability of the merely sentient, it begins to seem questionable whether the idea of respect for persons as ends in themselves can continue to play any useful role in ethics, and doesn't rather serve to obscure the most important issues.

2.3.3 The Problem of Objectification

The third problem that Code's discussion raises for the principle of respect for persons is slightly more difficult to capture. The first problem – that of integration – concerned the features of persons that require respect: how well-integrated they may or may not be, and, accordingly, to what extent it is possible to satisfy the requirements of beneficence etc., and respect for self-determination, under the unified umbrella of a principle of respect for persons. The second problem – that of fragmentation – concerned the hallmarks of personhood, and whether it is possible to pick out a distinctive morally significant class of persons at all. (And clearly these two problems are interrelated, for if the morally significant features of persons are not well-integrated we should probably anticipate difficulties in picking out a distinctively morally significant class of persons, all and only the members of which share all the relevant features.) The third problem – the problem of objectification – arises from a particular feature which ontological persons share, that serves to massively complicate the question of what is involved in respecting ontological persons *as* persons.

According to Code, individuals whose capacity for self-determination is fragile or intermittent face a particular danger in situations of professional care. This is the

danger of 'objectification'.[66] They are exposed to this danger because they typically do not have the resources to stave off attempts by care workers to *stereotype* them. By 'stereotyping' is meant the practice of considering them uncritically in the light of some generic notion of the type of individual they are. (For example, as nothing more than a 'poor thing'.) Code equates such stereotyping and objectification directly: 'Stereotyping people amounts to objectifying them', and objectification involves 'treating them as neatly classifiable items in the world'.[67] She goes on to present the objectification of clients by health care professionals as an ever-present danger in institutionalised settings.[68] At the same time however, the objectification of patients in such settings clearly also involves their instrumentalisation – their being treated as nothing more than a means to another's ends. Thus Code's conception of objectification incorporates both instrumentalisation and stereotyping, in a distinctive combination. Objectification involves treating persons as if they were no more than objects, things. Code observes:

> By now it is a matter of course to affirm without argument, at least in feminist philosophical circles, that persons are significantly different from objects, both epistemologically and morally, and that our attitudes to them must, accordingly, be quite different... Treating a person as little more than an object, in its cognitive dimension, implies acknowledging no significant differences between such a person and an ordinary, everyday object. Plainly, at the simplest level, there is a mistake here: one ought to know better. But the cognitive problem is more complex than this rather naive statement suggests. To acknowledge what, for want of a better term, I shall call the 'personhood' of another human being involves recognising responses, conditions, actions of persons as worthy of respect just because they are manifestations of the choices... of active, sentient, thinking beings.[69]

We have already considered some of the reasons why distinguishing persons from objects may not always be a straightforward matter. However, Code's discussion suggests a further danger: not simply that of *overlooking* another's capacity for some form of self-determination, but that of *actively undermining* their personhood. Ordinarily it is our capacity for self-determination that enables us to resist other people's attempts to stereotype us. And it seems that it is particularly where this capacity begins to fade that objectification becomes a threat. In order to appreciate the full significance of the point here, it is important to recognise that neither the dignity of persons, nor the *suggestibility* of persons, is linked with their capacity for effective self-determination in any straightforwardly linear manner.

If we think of dignity as proportional to our capacity for meaningful self-determination, we will tend to think that as the difficulty of distinguishing the genuinely self-determining from the non-self-determining increases, the moral stakes diminish. If self-determining agents possess dignity in the precise degree to which they possess a capacity for meaningful self-determination, then it seems

---

[66]Code 1995, pp. 83–4.

[67]Ibid., p. 87.

[68]Ibid., pp. 83–4.

[69]Ibid., p. 88.

that insofar as an individual's capacity for self-determination is wavering or intermittent, her dignity will also be wavering or intermittent. It will then seem that the challenge in such cases is simply that of identifying the cut-off point for meaningful self-determination with sufficient accuracy, so as to avoid potentially misclassifying persons as non-persons (or vice versa). But it is clear from Kant's discussion of the nature of the respect due to persons that dignity is not possessed in degrees. For Kant, all rationally self-determining beings possess the full measure of human dignity.[70] If he is right about this, then the moral stakes do not diminish as the difficulty of distinguishing the genuinely self-determining from the non-self-determining increases; and this greatly intensifies the associated moral challenge.

One thing that does diminish proportionally with the decline of effective self-determination though is the capacity to defend our own autonomous choices. Thus not only are the stakes as high as ever in such cases, but it is much easier for carers to finesse the moral challenges posed by declining self-determination by erecting artificial divisions between the self-determining and the non-self-determining. Under the pressure of professional imperatives, the quiet but vital signs of residual autonomy will be in danger of being swamped by institutional strategies designed to simplify and disambiguate an essentially difficult and sensitive situation.

Moreover, as the capacity for robust self-determination declines, so our associated capacity for self-interpretation becomes something of a Trojan horse. Those who are borderline, or intermittently-self-determining, will retain some capacity for self-interpretation even as their capacity for meaningful self-determination declines, and as a consequence they will become increasingly *suggestible*. In such a condition they will be exceptionally vulnerable to the objectifying attitudes of others; and increasingly prone to adopting the interpretations and evaluations suggested to them in an uncritical manner. As a result, just because they retain sufficient capacity for self-interpretation to actively collude with professional and institutional strategies that demand a neat and tidy dividing line between those who may be trusted to make decisions on their own behalf and those who may not, they will be in danger of effectively conspiring in their own reduction to the status of an object. Just where it is most vital that human dignity be protected and respected then – in those who still possess a flicker of autonomy and capacity for self-interpretation/self-determination – it is likely to be pressed into the service of its own denial. Residual personhood may then find itself not simply ignored, but actively undermined. The patient's residual capacity for self-interpretation becomes an enemy within.

Thus the full extent of the cognitive challenge in cases of potential objectification can only be appreciated by recognising that the moral stakes do not diminish as the difficulty of distinguishing those who have the capacity for self-determination from those who do not increases, and that it is precisely in those cases where it is most crucial that an individual's residual capacity for self-determination be acknowledged that it is most likely that it will be undermined from within. The way we treat vulnerable individuals influences whether they are able to understand

---

[70]Kant 1996, pp. 83–5.

themselves as self-determining, and consequently whether they are able to behave in significantly self-determining ways. The issue of who is able to qualify as a person is not independent of the question of who we are in practice prepared to treat as a person. The danger that vulnerable self-determining patients, in particular, will find themselves 'objectified' suggests that not only is the category of persons problematic in itself, but that self-determining personhood is highly susceptible to external influences.

2.4 The Aftermath

In the face of the various problems detailed above, it is appropriate to consider what the fate of the principle of respect for persons in contemporary applied and professional ethics can now be.

Kant's formula of humanity was evidently understood by its author to apply to our treatment of all and only humans. But we have seen that there is good reason to think that human beings are not the only morally considerable beings. For one thing, humans seem not to be the only significantly self-determining beings. And if it should be claimed in response that only the sort of sophisticated self-determination we typically find in human beings (O'Neill's principled autonomy, as opposed to individual autonomy), deserves to be regarded as a marker for personhood, we will find ourselves having to explain the relevance of this sort of sophisticated self-determination to moral patiency (in the face of the potential exclusion from the set of moral patients of those humans who are unable to achieve more than individual autonomy). Without a clear, relevant and plausible set of criteria for personhood then, which does not reduce to a set of markers for biological humanness, the principle cannot be consistently deployed.

Downie and Telfer's focus on personality rather than rational self-determination as such may seem to be a more promising approach. But if personality commands the various forms of respect associated with duties of beneficence as well as duties of respect for self-determination, it is again reasonable to ask what makes personality as such the unit of prime moral significance. Doesn't it simply represent the chance coming together of a range of features (sentience, a capacity for self-determination, a capacity for rational thought and action) that can be found scattered across the human and non-human worlds? If these features command our respectful consideration wherever they are found, then what work is the idea that they are sometimes all found together in association with something we call 'personality' able to do? (Other than underpin the sort of speciesist prejudices that Singer, Regan and others provide cogent arguments for doing away with.) In the face of these considerations, it is hard to resist the conclusion that the idea of the *person*, as the unit of prime moral significance, is an illusion we would be better off without.

A final recourse for adherents of the ethics of respect for persons might be to argue that the concept of the person was never intended to do any really important work, but was simply a convenient label for the locus of a range of different moral

duties, which we can continue to take seriously, regardless of whether we are able to identify a special group of individuals to whom they uniquely apply. This line of defence might seize on something like the four principles of Beauchamp and Childress as an example of an ethical approach that operates in practice as an ethic of respect for persons, without the encumbrance of a theoretical account of personhood. But can it really be said that an approach such as this, which deliberately dispenses with any elaborate notion of personhood, still qualifies as a 'respect for persons' ethic?

While the absence of any foundation in a coherent concept of personhood may initially seem, in the light of the problems we have considered, to be a strength of such an approach, its relative lack of theoretical encumbrances cannot be presented as an unalloyed advantage. Suppose we set out to put the four principles approach into practice, and abide by the principles of beneficence, non-maleficence, justice and respect for autonomy in our professional lives. If the approach also embodied an overarching principle of respect for persons, and a usable conception of personhood, we would presumably have a clear idea of where and when to apply these principles. But the approach aims to dispense with these encumbrances. How then are we to apply the principles in practice? If we are not to have further principles for applying the principles, the only conclusion seems to be that their application to any actual case must be considered to be self-evident, once we have undertaken the necessary specification and balancing (as, for example, in the case of a patient in a persistent vegetative state, the principle of non-maleficence would remain applicable, but that of respect for autonomy would not). Rather than insisting on applying all four principles to all and only those cases involving 'persons' then (which would entangle us in the various difficulties discussed above) we must ask in each individual case whether we can meaningfully apply each of the four principles, and, if the answer is affirmative, we should go ahead and do so. But if that is how the approach is supposed to work, then nothing but speciesism (and professional prudence) seems to prevent a veterinary doctor from applying the four principles approach to the full range of patients with which she deals. It is evident that the principles of beneficence, non-maleficence and justice could be as *meaningfully* deployed in veterinary medicine as in human medicine.[71] It does not seem quite so straightforward to apply the principle of respect for autonomy in a veterinary context, but even that seems possible. As we have seen, the 'autonomy' that Beauchamp and Childress have in mind is not full blown Kantian 'principled' autonomy, but what O'Neill terms 'individual' autonomy – which amounts to a capacity for making meaningful choices. But this is a capacity that Tom Regan claims to find spread well beyond the confines of the human world, and present throughout the mammalian world, from which the

---

[71] Of course, doing so might not always result in similar decisions being made – for example, where the principle of beneficence could be meaningfully employed, but that of respect for autonomy could not, it seems much more likely that we will find ourselves favouring euthanasia, in cases of serious incurable disease.

veterinarian's patients are likely to be predominantly drawn.[72] Thus it may not be *meaningless* to apply the principle of respect for autonomy in veterinary medicine.

Perhaps it will be objected that the principlist approach is devised for the context of professional ethics, and in particular for human healthcare ethics, and that any apparent problems generated by applying the approach to e.g. veterinary medicine can therefore be ignored, on the basis that they reflect an arbitrary extension of the approach beyond its proper sphere. But this won't really do. Ethics is not simply a matter of legislation, and if an ethical principle seems applicable beyond the sphere its originators had in mind, this is not something we can afford to dismiss. A number of social advances have come about through the extension of ethical and political principles beyond the sphere that the originators of those principles envisaged. And in these cases it is the attempt to restrict the scope of the relevant principles, rather than to extend them, that now looks arbitrary. It cannot be seriously argued then that we should not consider what would be involved in applying at least some of the four principles in dealing with non-humans (and it would have to be borne in mind in doing so that that we cannot always apply all four principles when dealing with human beings). This being so, the idea that the four principles approach will do duty as a kind of contemporary ethic of respect for persons, freed of inconvenient theoretical commitments to a particular concept of personhood, is evidently mistaken. While perhaps initially appearing to underwrite what Downie and Telfer characterise as 'the ordinary rules and judgements of social morality', the four principles approach seems to mark the death of any distinctive respect for persons ethic, rather than its salvation. The image of the veterinary doctor conscientiously abiding by the four principles vividly illustrates how the category of 'persons' no longer seems to mark any particularly significant ethical boundary.

Approaches such as that of Beauchamp and Childress seem to reveal, in their studied superficiality, a deep contemporary anxiety surrounding the ethics of respect for persons. Evidently, if an ethic of respect for persons is to survive, new philosophical arguments for the special moral significance of persons are required. At a minimum, these arguments will have to show that something closely analogous to a traditional principle of respect for persons is able to play a distinctive and significant role in contemporary ethics. They will need to show that some version of the principle of respect for persons captures important features of our ethical responsibilities to the sorts of beings that have traditionally been distinguished as 'persons' in a manner that (e.g.) the principle of respect for autonomy fails to do. The remainder of this study aims to do just that.

---

[72]Consider the case of a vet working in a research laboratory. It is simple bad faith to pretend that we cannot possibly know how animals feel about certain research procedures. As Steven Sapontzis observes (cited in Rollin 1998), it is easy enough to find out what they would prefer – open the cages.

References

Aristotle. 1995. *Ethics*. London: Penguin.

Beauchamp, Tom L. 1999. The failure of theories of personhood. *Kennedy Institute of Ethics Journal* 9:309–324.

Beauchamp, Tom L., and James F. Childress. 2001. *Principles of biomedical ethics*, 5th edition. Oxford: Oxford University Press.

British Medical Association. 1993. *Medical ethics today*. London: BMA Publishing.

Carne, Ros. 2010. *Professional ethics* (Bar Manuals). Oxford: Oxford University Press.

Code, Lorraine. 1995. Persons, and others. In *Rhetorical spaces*, 83–102. London: Routledge.

Descartes, Rene. 1986. *Meditations on first philosophy* (trans: Cottingham, J.). Cambridge: Cambridge University Press.

Dillon, Robin. S. 1992. Care and respect. In *Explorations in feminist ethics*, eds. Eve Browning Cole and Susan Coultrap-McQuin, 69–81. Bloomington: Indiana University Press.

Downie, R.S., and Elizabeth Telfer. 1969. *Respect for persons*. London: Allen and Unwin.

Hume, David. 1969. *A treatise of human nature*, ed. Ernest C. Mossner. London: Penguin.

Kant, Immanuel. 1929. *Critique of pure reason* (trans: Kemp-Smith, N.). London: Macmillan.

Kant, Immanuel. 1996. Groundwork of the metaphysics of morals. In *Immanuel Kant: practical philosophy*, ed. Mary Gregor, 41–108. Cambridge: Cambridge University Press.

Locke, John. 1924. *Two treatises of civil government*. London: Dent Dutton.

Macintyre, Alasdair. 1981. *After virtue*. London: Duckworth.

Mill, John Stuart. 1962. Utilitarianism. In *Utilitarianism*, ed. Mary Warnock, 251–321. London: Fontana.

Nell, Onora. 1975. *Acting on principle*. New York, NY: Columbia University Press.

O'Neill, Onora. 2002. *Autonomy and trust in bioethics*. Cambridge: Cambridge University Press.

Plato. 1970. *Laws* (trans: Saunders, T.J.). London: Penguin.

Regan, Tom. 1984. *The case for animal rights*. London: Routledge & Kegan Paul.

Regan, Tom. 1985. The case for animal rights. In *In defence of animals*, ed. Peter Singer, 13–26. Oxford: Blackwell.

Regan, Tom. 2003. *Animal rights, human wrongs*. Lanham, MD: Rowman & Littlefield.

Rollin, Bernard E. 1998. The moral status of animals and their use as experimental subjects. In *A companion to bioethics*, eds. Helga Kuhse and Peter Singer, 411–422. Oxford: Blackwell.

Singer, Peter. 1993. *Practical ethics*. Cambridge: Cambridge University Press.

Singer, Peter. 1995. *Animal liberation*. London: Pimlico.

Singer, Peter. 2002. *Unsanctifying human life: Essays on ethics*. Oxford: Blackwell.

Tooley, Michael. 1998. Personhood. In *A companion to bioethics*, eds. Helga Kuhse and Peter Singer, 117–126. Oxford: Blackwell.

Wood, Allen W. 1999. *Kant's ethical thought*. Cambridge: Cambridge University Press.

Chapter 3
Discrimination

Abstract Given the fragmentation of the idea of the moral person described in Chapter 2, one may wonder whether the principle of respect for persons can have any relevance to contemporary ethics. I will argue that the apparent irrelevance of the principle is an illusion, stemming from an excessively restrictive view of the phenomena with which ethics must deal. This chapter begins the process of demonstrating the ongoing relevance of the principle, focusing on the phenomenon of discrimination. Discrimination proves to be more than just procedural injustice, since it crucially involves the interpretation of its victims. To be discriminated against is to be treated in a procedurally unjust manner *as* a member of this or that social group. Discrimination thus proves to be an *interpretive* moral wrong. The existence and nature of interpretive moral wrongs suggest that moral patiency is a more complex matter than might have been thought, on the basis of traditional accounts of respect for persons.

3.1 Discrimination and Procedural Unfairness

Given the multiple fragmentation of the idea of the moral person as an end in herself described in Chapter 2, one may well wonder whether the principle of respect for persons can have any relevance to contemporary applied and professional ethics. Both the fact that the concept of a person is now so fragmented, and the fact that key higher-level principles, which might once have been thought to be derivative of a principle of respect for persons (such as the principle of justice), are widely held to stand in no need of support from such a foundational principle, may suggest that there is no significant role for a principle of respect for persons to play. Nevertheless, in the remainder of this book, I want to make a case for the continuing ethical relevance of the principle of respect for persons, in a revised and re-interpreted form. My aim will be to show that the apparent contemporary irrelevance of the principle of respect for persons is an illusion, stemming from an excessively restrictive conception of the ethically relevant features of ontological persons, and an equally restrictive view of the range of phenomena with which ethics must deal.

The first stage of the process, to be undertaken in this chapter, will be to introduce a class of distinctive moral phenomena, which I will refer to as *interpretive* moral wrongs. The existence and nature of these wrongs suggest that moral patiency is a more complex matter than might have been thought, on the basis of traditional

accounts of respect for persons. I will go on (Chapters 5 and 6) to explore the ways in which interpretive moral wrongs pose particular threats to ontological persons, and I will argue that it is in order to fully protect ontological persons that an ethical principle of respect for persons is still very much required. My first example of an interpretive moral wrong is that of discrimination.

At the mention of the word 'discrimination' it is likely that most people will think first of the sorts of informal practices in employment and social welfare provision etc., that serve to ensure that certain independently identifiable ethnic, religious and gender groups are systematically disadvantaged, in a more or less deliberate and calculating manner. On the basis of these examples, it might well be thought that the offence of discrimination is reducible to a form of social injustice – paradigmatically, to a distributive injustice. On this sort of view, one would be discriminated against insofar as one was the victim of a process designed to unfairly apportion certain social goods or harms. But while many discriminatory procedures and institutions take this form, not all examples of discrimination conform to the model of distributive injustice.

Systems such as the former race laws of the southern United States, the former apartheid system of South Africa, and the former structurally sectarian legal and political system of Northern Ireland, are paradigm cases of discriminatory social institutions. And these institutions, as the experiences of the twentieth-century civil rights movements that opposed them attest, were not reducible to engines of distributive injustice. Participants in those movements were often campaigning in part for a fairer distribution of social goods and harms, but they were also campaigning for the recognition of basic rights and freedoms – such as the freedom to participate fully in democratic political processes. To be denied the right to political participation on grounds of gender, ethnicity or religious affiliation is clearly discriminatory. But this denial represents a denial of basic rights, rather than (simply) a denial of legitimate interest claims.

Discrimination is not reducible to distributive injustice then. Moreover, even if we focus on those cases of discrimination that do involve distributive injustice, there turns out to be more to discriminatory treatment than distributive *unfairness*. Lotteries, for example, typically distribute goods in a distinctly unfair manner. But we can hardly claim that the distribution resulting from a lottery is unjust, so long as the lottery is properly run.[1]

Showing that discriminatory distributive injustice is not reducible to simple distributive injustice is not however as straightforward as showing that distributive injustice is not reducible to unfairness. Before confronting the question of the relationship between discriminatory distributive injustice and simple distributive

---

[1]In fact this involves a considerable oversimplification: faced with an existing unjust distribution of goods it may well be the case that a properly run lottery perpetuates and exacerbates pre-existing social injustices, through e.g. a combination of ticket price, unfair distribution of profits, supporting socially divisive patterns of aspiration etc. The main point however stands: if some conceivable distributively unfair institutions are not unjust – lotteries, in ideal circumstances – then distributive injustice is not reducible to distributive unfairness.

injustice directly, it will be helpful to pursue the comparison with a properly run lottery further, as an instance of distributive unfairness that is not (ideally) unjust.

The case for the claim that a properly run lottery doesn't necessarily result in an unjust distribution of goods, despite the fact that that it inevitably results in an unfair distribution, rests on the idea that while lotteries typically result in unfair distributions of goods, they are not, when properly run, *procedurally* unfair. There will be winners and losers in any lottery. And any loser is likely to feel that the result is unfair – as indeed it may well be: seldom if ever are we likely to be in a position to say of a lottery winner that she deserved her success. The loser's complaints are only likely to carry any weight though if it is possible to point to some feature of the running of the lottery that unfairly weighted the outcome in favour of the eventual winner. In a properly run lottery everyone competes on equal terms. And in any case in which it is apparent that players were not in fact competing on equal terms, it seems reasonable to say that the lottery outcome is not simply unfair, but positively unjust.

In respect of lottery-type processes at least then, it seems clear that unfairness only implies injustice when it takes the form of procedural unfairness, such that players do not compete on equal terms. That said, it is important to note that procedural unfairness may take more than one form. Let us distinguish two types of procedural unfairness: 'loading' unfairness, and 'rigging' unfairness.[2]

Where loading-type unfairness is present, players do not compete on equal terms, because the 'dice' are effectively loaded in favour of certain players, and/or loaded against certain others. Imagine, as an example of this type of unfairness, the case of a badly-run lottery in which the bulk of the tickets bought in certain districts are effectively duds, since as a consequence of (avoidable) delivery and communication problems, they are not represented in the final draw at all. In this case the 'dice' are effectively loaded. Those who bought tickets in the affected districts are systematically disadvantaged, by comparison with other players.

In the face of protests concerning the injustice of the result of such a lottery one can imagine the lottery organiser responding that the original result ought to stand. There was no deliberate unfairness involved, since it has to be considered a matter of sheer chance that those tickets that found their way into the final draw got there at all. The result cannot be considered unjust, the organiser might continue, because all that has really happened is that the element of randomness that would otherwise have been confined to the draw itself has leaked out, and contaminated an earlier stage of the process. Formally, the organiser may seem to have a point here. If nobody was actually cheating, and the result was still effectively random, then it is true that certain possible forms of procedural unfairness have been avoided. But it seems unlikely that any of the unlucky players would be satisfied with this. Random or not, the result of the lottery was more than simply randomly unfair (as we must expect any lottery to be), because the players were not, as they should

---

[2]I do not mean to suggest that this represents an exhaustive classification of forms of procedural unfairness.

have been, competing on equal terms. The ticket holders were not equally affected by the procedural irregularities that occurred. This sort of case seems to be one in which we can reasonably say that the unfair distribution resulting from the lottery draw was linked to procedural unfairness in such a manner as to make the resulting distribution of prizes not simply unfair – as would have been the case even in a properly run lottery – but positively unjust. And we would certainly tend to feel that disappointed players would be justified in demanding that the lottery either be re-run properly, or the ticket price refunded.

With this illustration of loading-type procedural unfairness in mind, we can illuminate rigging-type procedural unfairness, by contrast with it. In the case of rigging-type unfairness we are no longer dealing with procedural irregularities which affect players more or less randomly, and which may result merely from incompetence, but with systematic irregularities, of a kind that might result from deliberate cheating. Imagine for example a parallel lottery case, in which it turned out that rather than being randomly absent from the final draw, the 'lost' tickets had been deliberately dumped by the lottery organiser, who had arranged for this to happen to all tickets being returned from geographical regions in which he had no family members, in order to improve the chances of a member of his own family winning.

In this case, as in the previous case, we would tend to consider the final result to be not simply unfair but positively unjust. But at the same time we would take a rather different view of the process, and of the conduct of the lottery organiser. In both types of case elements of procedural unfairness serve to transform what might otherwise have been a merely unfair distribution into a positively unjust distribution. But the types of injustice involved in the two cases are significantly different. In a properly run lottery players compete on equal terms, and the inevitably unfair result cannot (in properly egalitarian circumstances at least) be said to be unjust. In a lottery that is run incompetently, but in good faith, the result may be not simply unfair, but positively unjust, since players do not compete on equal terms. But in a dishonestly run lottery, while the result is again unjust, the injustice involved is of a rather different character. Here some of the players are not simply systematically disadvantaged, they are the victims of cheating.

What, then, lies at the basis of the distinction between loading-type and rigging-type procedural unfairness? The main thing we will look for when trying to ascertain whether we are dealing with a case of rigging, rather than mere loading, is, as has already been suggested, evidence of the sort of procedural irregularity that indicates intent to systematically advantage certain individuals or groups, and/or systematically disadvantage other individuals or groups. Loading-type procedural unfairness may be accidental, and simply unfortunate. But rigging-type unfairness is always in some sense or to some degree deliberate.

A further distinguishing feature of rigging-type unfairness derives from this fact that rigging is always to some degree deliberate. In the case of loading-type unfairness there is an evident logical symmetry in the way that a procedure loaded against one individual or group is loaded in favour of others. Any procedure that is loaded in favour of one individual or group will automatically and without further ado be loaded against all other players. By contrast, a rigged process may, but need not,

display the same sort of symmetry. A process may be rigged in favour of an individual or group as its beneficiary, and/or rigged against some other individual or group as its victim. But from the fact that an individual or group is the victim (or beneficiary) of rigging it does not follow that that individual or group is the *intended* victim (or beneficiary) of rigging. Thus it does not automatically follow that the process is rigged *against* (or in favour of) them purely because it is rigged in favour of (or against) another individual or group – though it will certainly be loaded against (or in favour of) them.

Given the nature of rigging then, it is possible to be *only incidentally* the victim (or beneficiary) of rigging. In the 'loading' case the conclusion that the process is loaded against someone follows straightforwardly from the fact that they are systematically disadvantaged, as compared with other players. But a rigged process may be intended to favour an individual or group, without being intended to disadvantage others.[3] In any case in which the victims (or beneficiaries) of rigging are not the intended victims (or beneficiaries), we cannot reasonably claim that the process was rigged in their favour (or against them), though we can still claim that it was loaded against them (or in their favour). In this sort of case we will have to say that they are the incidental victims (or beneficiaries) of a process that was rigged in favour of (or against) someone else. The possibility that one might be the incidental victim or beneficiary of rigging has no parallel in the case of loading, since the absence of the sort of intent that marks out a given case of procedural unfairness as a case of rigging has the consequence that everyone affected by loading is affected on the same basis.

It is important to note at this point that the question whether a given case of rigging does in fact display the sort of symmetry that we find in cases of loading is heavily context-dependent. In the second of the lottery examples discussed above – the one used to introduce rigging-type unfairness – it might well be plausible for the organiser to claim, in mitigation, that while the lottery was certainly rigged to favour his own family members the process was not *rigged against* anyone (though it was certainly *loaded* against players buying tickets in areas where the organiser happened not to have family). In this sort of case, assuming the lottery itself was a suitably large-scale affair, conducted over a wide geographical area, the organiser's claim may be reasonable – at least to the extent that he can plausibly plead ignorance of those adversely affected. Of course, in any case in which there can only be a limited number of winners it will follow that anyone not advantaged by the relevant procedural irregularities will be disadvantaged by them. But it may still be plausible for the organiser to claim that although these faceless (to him) others were certainly disadvantaged by his actions, this fact never entered into his calculations. These unfortunate players were only incidentally the victims of rigging. By contrast, in a case where it is similarly impossible to advantage one player without

---

[3]It is of course perfectly possible for a rigged process to be intended to favour one group, and simultaneously intended to disadvantage others, or indeed for it to be designed simply to disadvantage one individual or group, with no thought at all being given to who might be thereby advantaged.

simultaneously disadvantaging others, and where the guilty party or parties cannot plausibly claim ignorance of those disadvantaged, and the effects of rigging upon them (say in the election of a small number of candidates to a committee, or in a job selection process), any instance of rigging will tend to display the kind of symmetry we observed in the case of loading. In any such instance we will tend to conclude that no one affected by the relevant act of rigging could have been affected 'only incidentally'.

3.2 Discrimination and Intentionality

What the foregoing discussion serves to bring out is that that rigging is significantly *intentional*. It is 'intentional' in two distinct but not unconnected senses. Firstly, rigging is intentional in the familiar sense of being deliberate, or quasi-deliberate. By 'deliberate' I mean that the perpetrator(s) had thought through the consequences of their action(s), and had taken a decision – tacitly or otherwise – to act as they did, so as to produce the relevant procedurally unfair result. My use of 'quasi-deliberate' probably requires a little more explanation. What I have in mind here are cases in which there would be a reasonable standing expectation that the perpetrators should have thought through the consequences of their actions, and acted accordingly. In other words, rigging-type procedural unfairness is quasi-deliberate in any context in which ignorance of our duties or of the consequences of our actions would, given our role and its responsibilities, be no excuse.

Secondly, rigging is intentional in the more specialised sense that that term has in the philosophy of mind. Here 'intentionality' refers to the referential and/or connotative quality of a mental state or act, such that it is 'of' or 'about' some independent object or state of affairs.[4] All intentional mental states/acts have a certain 'aboutness' or 'directedness-toward'.[5] So, for example, while both my toothache and my belief that my dentist enjoys inflicting pain are equally mental states, toothache is not an intentional mental state, since it is not 'about' anything, whereas the aforementioned belief is an intentional mental state, because it is 'about' or 'directed toward' (refers to) my dentist.

The modern understanding of intentionality, and in particular the idea that intentionality is the hallmark of the mental is largely due to Edmund Husserl, the founder of the phenomenological school.[6] However, Husserl's pupil, Martin

---

[4]The term originates in medieval philosophy, but its modern usage derives from the work of Brentano, who in his *Psychology from an Empirical Standpoint* observes: 'Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object... Every mental phenomenon includes something as object within itself... In presentation something is presented, in judgement something is affirmed or denied, in love loved, in hate hated, in desire desired and so on.' (Brentano 1995, p. 88.)

[5]Note that as the above quotation makes clear, for Brentano (and often, as a consequence, for those influenced by him), all mental acts are intentional.

[6]See Husserl 1982, sections 36, 37.

Heidegger, challenged and radically modified Husserl's concept of intentionality. For Heidegger, the paradigm case of an intentional relation is not (as it is for Husserl), that between a mental act and some ideal mental content – such as a geometrical object, e.g. an ideal triangle – but between an acting human being and the object of his/her actions – which can be such a thing as a window, or a railway station, for instance.[7] Our relationship with the railway station is an intentional one in cases where our conduct toward it picks it out and refers to it in a manner that signals some understanding of it – as when we leave our seat and approach the door of the carriage as the train pulls into the station. In such a case we need not be assumed to be entertaining explicit thoughts about the railway station, let alone entertaining beliefs about it, for it to be an intentional object for us. On the Heideggerian conception of intentionality then, intentionality is a feature of our basic 'comportment' toward objects, rather than simply of purely mental acts (though of course it can still be a feature of mental acts).

As we have seen, a lottery or similar distributional procedure is never simply rigged. It is rigged against this or that group or individual, or in favour of this or that group or individual. Logically then, one cannot claim that a procedure is rigged without implying the existence of an independently identifiable individual or group with respect to whom it is rigged. The process of rigging *intends* or refers to its victims (or beneficiaries) in a manner analogous to the manner in which the mental act of thinking that my dentist enjoys inflicting pain 'intends' the man in question – since without an intended victim or beneficiary there is no act of rigging. By contrast one can perfectly well speak of a procedure being loaded without thereby implying that the dice are loaded against or in favour of any independently identifiable individual or group. It is because rigging is intentional in this sense that it displays the logical asymmetry that is not a necessary feature of loading-type unfairness. In the example of the rigged lottery discussed above, the lottery was rigged in favour of the organiser's family. In such a case it is possible to say that, in a purely formal sense, the lottery was rigged against the set of non-members of the organiser's family. But this would not make it implausible for the lottery organiser to insist that the process was not in fact rigged *against* anyone. After all, the set of 'non-members of the organiser's family' doesn't really seem to be the sort of independently identifiable group that any procedure *could* be geared to disadvantaging. The process was certainly loaded against the members of this group, and they were victims of a rigged process, but that process was not rigged *against* them.

The intentional quality of acts of rigging, which serves to distinguish them from cases of loading, suggests that the plausibility of a charge of rigging is conceptually tied to the independent identifiability of its victims and/or beneficiaries. One needs to be, from the standpoint of the alleged perpetrator, significantly identifiable – as this or that individual, as a member of this or that group – to be a potential victim of rigging. One needs to have the kind of independent identifiability that could serve as an adequate attachment-point for the intentionally-structured act of rigging.

---

[7]See Heidegger 1988, sections 9b, 9c.

It is however important to reiterate an earlier qualification at this point. The asymmetry I am claiming for rigging is to a degree context-dependent. While in the lottery cases discussed above it is plausible to say that, given the relative anonymity of the whole process, certain victims or beneficiaries will be victims or beneficiaries only incidentally, there are other cases in which this will not be so. In a job selection process for instance, where the numbers are relatively small, and the selection panel has detailed information on each candidate, it will usually be implausible to suggest that the process might be rigged in favour of certain candidates but not at the same time rigged against the others. In such a case the panel cannot plead ignorance, or the disinterest that comes with the various forms of distance that might otherwise interpose between victims and perpetrators. And even in a case in which an individual can honestly claim to have been ignorant of the fact that rigging the process in favour of a preferred candidate effectively amounted to rigging the process against the others, we will tend to take the view that those concerned ought to have been aware of the likely consequences of their actions, and that therefore the rigging was in the relevant senses intentional, being both deliberate (or quasi-deliberate, in the sense discussed above) and occurring in a context in which the perpetrator had a duty to take the impact on the disadvantaged candidates into account.

3.3 Discrimination as an Interpretive Moral Wrong

With the above points in mind, we can return to the main topic, that of discrimination. Recognising that rigging-type procedural unfairness is intentional in both of the above senses is key to understanding the nature of discrimination. Sticking for the moment with the case of distributive procedures, I wish to claim that a distributive procedure is discriminatory if it is not simply unfair, but is unjust, and if the injustice in question is such that the procedure can be said to have been rigged against its victims.

At this point however it is necessary to introduce one further refinement. Nothing in the characterisation of discrimination just given rules out the possibility that individuals might be victims of discrimination, just as groups are. We would though, I think, reject the idea that discrimination could occur in a case in which rigging was in evidence, but the intentional character of the relevant act of rigging was such that it picked out, and only picked out, a specific individual as the intended victim. It is evident from the lottery examples discussed above that both individuals and groups can be the intended victims (or beneficiaries) of rigging. But it seems necessary that when an individual suffers from discrimination she does so 'as' a member of a group. One may suffer injustice, as a result of rigging, both as a private individual, and as a member of a group. But one may only suffer discrimination 'as' e.g. a Jew, a woman, a person of mixed race, etc. (And when an individual suffers as a result of an act of rigging we will require evidence that she suffered that injustice 'as' a member of some independently identifiable group before we conclude that she has been the victim of discrimination.)

The difference between injustice suffered by an individual as an individual, and injustice suffered by that same individual as a member of a group seems to lie in the specific intentional character of discriminatory rigging. The process of rigging must, as we have seen, be intentional in the sense of being deliberate, or quasi-deliberate – the perpetrator must either have understood what he was doing, or have been acting in some capacity (for example in a professional role) such that he had a duty to be aware of the impact of his actions on their victims/beneficiaries. In order to pick out what makes an instance of rigging-type unfairness authentically discriminatory though, we need to draw a further distinction in respect of intentionality in the sense of 'aboutness'. I have said that mental acts etc. may be said to be intentional in the sense of referring directly to this or that individual person or thing (e.g. my dentist). Alternatively however, they may be said to be intentional in the sense of having a specific content. A belief that purports to be true of all dentists, for example, will be 'of' or 'about' dentists in general. In this case the belief will certainly refer to individual dentists. But it will refer to them (relatively) indirectly. It will refer to individual dentists simply qua dentists, rather than qua the particular individuals they are. The intentional quality of such a belief is not dependent on its referring to any particular dentist, living or dead (or imaginary). By contrast one cannot be said to have a mental image of a particular dentist unless that image refers directly to the individual in question. Quite how it is that mental acts come to possess this mysterious quality of intentionality is a philosophically controversial question, which thankfully we are under no obligation to pursue further here.[8] It is sufficient for present purposes to note that that a mental act etc. may possess a specific quality of intentionality such that it is of or about specific individuals (living, dead or imaginary), and only incidentally about the groups to which they belong, or it may possess a specific quality of intentionality such that it is of or about a group,

---

[8]Quite how mysterious is underlined by Hilary Putnam's rather hostile discussion of the phenomenological doctrine of intentionality, in Putnam 1981, chapter 1. According to Putnam, the doctrine amounts to a 'magical' theory of reference – mental images must 'magically' reach out to objects in the world. But Putnam misinterprets the phenomenological approach in two crucial ways. Firstly, the phenomenologists are concerned with conscious *acts*, rather than with mental or physical *images*. While the claim that a mental *act* intrinsically refers to its extra-mental object may perhaps also strike Putnam as implausible, its implausibility must surely have some other basis – unless we are to believe that there are no relevant differences between mental acts and physical images (Putnam's example) in this regard. Secondly, the phenomenologists treat the phenomenon of intentionality (for Putnam, equivalent to 'reference') as fundamental. Their doctrine of intentionality cannot legitimately be presented as a false theory of reference then, because is it not put forward as an explanatory *theory* of reference at all. On the other hand, the Wittgenstein-derived alternative Putnam prefers does seek to provide an explanation of reference. And as such it is plainly inadequate in many respects (how, for example, is it possible to explain the capacity of a dream to refer, at the very time at which it is unfolding, purely on the basis of social and linguistic practices?). Note, finally, that my claim that some moral wrongs display an intentional element does not ultimately stand or fall with the correctness of the phenomenological view of intentionality. In this case at least, the more we tend to think of the referential quality of an act as a function of social practices, rather than of private mental states, the more plausible will the claim that some wrongs 'refer' to their victims appear to be.

and only incidentally about its members. The differences in the intentional features
of each type of act are not however reducible to the *directness* with which they refer
to individuals. For when I refer directly to an individual, I *simply* refer to her, I
do not refer to her *as* anything. My belief that 'S. W. enjoys inflicting pain' refers
to my dentist. But it does not refer to him qua dentist. On the other hand, when
I refer indirectly to an individual, via an act that refers to the group to which she
belongs, I necessarily refer to her as a member of the group in question. If I believe
that dentists in general enjoy inflicting pain, this belief refers indirectly to S. W.,
and refers to him *as* a dentist (that is, in his capacity of being a dentist). Earlier we
saw that a rigged process 'intends' those it is rigged against. But a rigged process
that intends a specific individual simply intends that individual. It need not, as is
now apparent, intend him or her *as* anything. A rigged process that constitutes an
instance of discrimination, by contrast, intends its victims *as* that which they are,
and necessarily does so indirectly. The victims of all acts of discrimination are, of
course, individuals. But the discriminatory act intends them only as members of
an independently identifiable group. To be the victim of discrimination then, is to
be the victim of an injustice which is rigged against its victims as members of an
independently identifiable group.

With this understanding of discriminatory processes in mind, it is a relatively
straightforward matter to argue that discrimination need not be confined to cases of
distributive injustice. As the example of discrimination in small-scale elections has
already indicated, a process or institution may display procedural unfairness of an
appropriate intentional sort to qualify as a case of discrimination without having any
connection with the distribution of social goods. If, from the standpoint of prospec-
tive candidates, holding political office can plausibly be presented as a social good,
which can be granted or withheld justly or unjustly, any such process that is rigged
against certain candidates as members of a particular group will be discriminatory.
At the same time, while the right to political representation is more appropriately
presented as a right than as a social good, the same electoral procedure will simul-
taneously be rigged against that candidate's natural constituency. And where the
relevant act of rigging displays an appropriately intentional character, we will tend
to say that the process discriminates against the relevant constituency.

The above analysis of discrimination has concentrated on examples of distribu-
tive injustice stemming from procedural unfairness. The conclusions drawn are
however intended to have more general application: I have tried to capture what
is central to all examples of discrimination. Before closing this stage of the dis-
cussion however, it is necessary to say more about the general applicability of the
conclusions drawn above.

My central examples have been of lotteries, and, of course, the similarities
between lotteries and more general mechanisms of social distribution are not always
very close. Questions of distributive justice do sometimes arise in lottery-type con-
texts. For example, lottery-type distributional processes have in some cases been
used in the allocation of scarce medical resources.[9] Our discussion of loading-type

---

[9]Beauchamp and Childress 2001, pp. 268–9.

procedural unfairness and rigging-type procedural unfairness has application to these cases, and suggests that there need not in principle be anything unjust about making distributional decisions on a lottery basis, despite the fact that the results must always be unfair. For example there is in principle no incompatibility between adhering to Beauchamp and Childress's principle of justice, and advocating the use of lottery-type processes to allocate scarce resources.[10] But while adhering to the principle of justice would not necessarily require us to oppose lottery-type distributional arrangements, it would require us to ensure that the relevant arrangements operate in a procedurally fair manner. Both loading-type and rigging-type unfairness would be equally ruled out by the principle of justice. And if the rigging in question intended members of an independently identifiable group as members of that group, then, on the foregoing analysis, we would have to consider it discriminatory.

Such examples of procedural unfairness in distributive contexts can be usefully extrapolated to cover cases that do not involve lottery-type decision-making (i.e. cases where we do not automatically anticipate an 'unfair' result), and the above discussion has touched on some of these cases, for example on job selection procedures, and small-scale elections. It is important to acknowledge though that not all mechanisms of social distribution involve the sorts of processes that can meaningfully be spoken of as *procedurally* either fair or unfair. Economic systems for example are usually more meaningfully characterised as structurally unfair (if unfair at all), than as procedurally unfair. Nevertheless the loading/rigging distinction remains applicable to them.

Market-based economic systems cannot, as a whole, be reasonably suspected of being procedurally unfair (though particular institutions within them certainly may be). Nevertheless the loading/rigging distinction may be applicable to a given market-based system, because the unfairness it displays, though structural rather than procedural, has a sufficient amount in common with the types of procedural unfairness we have been discussing. There exists a good deal of evidence to show that, as a result of fairly well-understood mechanisms, some market-based systems tend to ensure that those born poor stay poor, while those born rich have a much better chance of ending up rich. The dice are effectively loaded against the poor under such systems, and this entitles us to claim that these systems are not merely unfair, but positively unjust. At the same time however, such systems *may* simply display loading-type unfairness, rather than rigging-type unfairness, since the fact that a particular individual belongs to some independently identifiable group (for example, a particular ethnic, religious or gender group), *may* not be relevant to his or her economic well-being, under the relevant system. That said, the foregoing analysis of forms of procedural unfairness, and their connection with discrimination, has given us a good idea of what sort of structural unfairness we would need to look for in order to determine whether or not a particular economic system was discriminatory. In a case where a particular system not only tended to ensure that those born poor stayed poor, but also tended to ensure that members of particular

---

[10]Ibid., p. 271.

independently identifiable ethnic, religious or gender groups were particularly likely to remain poor if born poor, we would have reason to consider it discriminatory. Though of course, speaking strictly, we would have to say that the 'rigging' that was in evidence was a result of structural unfairness rather than of procedural unfairness.

A distinctive feature of cases of this type is that they raise questions about what I have called the 'intentionality' of discrimination, specifically, in respect of its deliberate or quasi-deliberate character. In cases of suspected *procedural* rigging, we would look for evidence of processes that are deliberately designed to favour or disadvantage certain individuals or groups. In the case of suspected *structural* rigging on the other hand, the assumption that a process that functions to advantage or disadvantage certain independently identifiable groups is designed, deliberately or quasi-deliberately, to do so, is by no means a safe one. It is possible to speak of both natural and social processes as operating 'as if designed' to have a particular result. In contemporary discussions, we would tend to view references to apparent design in nature as poetic at best. Thus we would tend not to attach any great significance to the observation that a given natural process operated 'as if designed' to bring about a particular result. With social processes however things are very different. Social processes range from the deliberately designed and heavily 'procedural' – for example systems of political representation – to the relatively organic and 'structural' – for example traditional patterns of land use. Confronted with this range, and faced with a system that functions so as to systematically disadvantage some independently identifiable group, it is evidently always going to be a complex matter to determine whether it does so deliberately (or quasi-deliberately), in a manner that entitles us to say that it is rigged against the group in question, and is thus discriminatory. Without in any way wishing to underplay the difficulty of this task, and without wishing to pre-empt decisions in any particular type of case, the important point for present purposes is that we do seem able to speak perfectly coherently of the intentionality of a social process that is structurally (rather than procedurally) rigged against members of an independently identifiable group, and is rigged against them quasi-deliberately *as* members of that group. The foregoing analysis would suggest that in such a case we would be dealing with a discriminatory social process.

Evidently, discrimination might be highly self-conscious in some cases, and ingrained and habitual in others. Nevertheless, we can say that a properly reflective agent who plays a key part in a discriminatory process ought to have been aware of what he was doing; and in such cases we will tend to find his actions vicious, even if he can honestly claim to have been acting without any great forethought, and without any particular goal in view. This may be especially important in professional-client relationships, where the existence of a professional duty of care may entitle us to find fault in cases where, given a different regulatory context, we might have been inclined to acquit the individual concerned of any wrongdoing. Acts of injustice whose intentional character is such that they pick out individual victims purely *as* individuals, are, as noted above, simply too personal to rank as discriminatory. If on the other hand the intentional character of a given systematically unjust process picks out its victims as members of an independently identifiable group, we have good reason to think we are dealing with a case of discrimination. It follows

that one cannot maintain a genuinely discriminatory practice except on the basis of some kind of explicit or implicit *ideology*.

It is the intentional and ultimately ideological quality of discrimination that makes it both peculiarly pernicious, and of particular interest for our analysis – since it begins to suggest that there is still an important and distinctive role for an ethic of respect for persons to play. Specifically, it suggests that the forms of respect due to persons as ends exceed those due to them simply as sentient beings, and as the possessors of 'individual' autonomy. Ensuring that social processes are procedurally and/or structurally fair will, on the basis of what has been said, ensure that they are not unjust. However, an analysis that focuses exclusively on justice as fairness will be insensitive to the difference between acts of simple injustice, and acts of discrimination, as outlined above. In this however discrimination proves to be only the tip of a moral iceberg. To fully appreciate the scale of the problem we need to pursue our enquiry further, initially in the area of a closely associated ethical phenomenon: that of stereotyping.

References

Beauchamp, Tom L., and James F. Childress. 2001. *Principles of biomedical ethics*, 5th edition. Cambridge: Cambridge University Press.

Brentano, Franz. 1995. *Psychology from an empirical standpoint* (trans: Rancurello, A.C. et al., 2nd edition). London: Routledge.

Heidegger, Martin. 1988. *The basic problems of phenomenology* (trans: Hofstadter, A.). Bloomington, IN: Indiana University Press.

Husserl, Edmund. 1982. *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy* (trans: Kersten, F.). Dordrecht: Kluwer.

Putnam, Hilary. 1981. *Reason, truth and history*. Cambridge: Cambridge University Press.

Chapter 4
Stereotyping

Abstract Chapter 3 argued that discrimination is distinguishable from other forms of injustice because it embodies, and puts to work, an interpretation of its victims. In discrimination individuals are systematically disadvantaged *as* members of independently identifiable groups. Nevertheless, there are contexts (such as selling insurance) in which it is considered permissible to make such judgements about individuals, which lead to their being systematically disadvantaged, but which we do not consider discriminatory. Consideration of these potential counterexamples highlights the role of *stereotyping* in discrimination. In addition to being a factor in discrimination, stereotyping proves to be an interpretive moral wrong in its own right. Stereotyped judgements made about individuals as members of identifiable social groups may be true or false, but are not morally innocent on the condition that they are true, since they may involve judging individuals on the basis of factors that affect them only as a matter of regrettable historical contingency.

4.1 A Potential Counterexample

The previous chapter sought to analyse discrimination as an interpretive moral wrong. When an individual suffers an injustice as the outcome of a process that is rigged against him or her as a member of an independently identifiable group she is discriminated against. Discrimination is 'intentional' both because it is deliberate, or quasi-deliberate, and because those who suffer discrimination do so 'as' members of such-and-such a group. In this latter respect there is a notable disanalogy between discrimination and other forms of injustice, since it is by no means necessary that individuals suffer other forms of injustice (e.g. economic exploitation) 'as' members of an independently identifiable group, or indeed 'as' the individuals they are, in any significant sense.

In order to fully appreciate the intentional character of discrimination however, it is important to consider the connection between discrimination and *stereotyping*. I have said that discriminatory distributive injustice can be distinguished from other forms of distributive injustice because in the former case an unjust process is intentionally structured. In many cases, perhaps the majority, the relevant intentional quality will derive from a particular set of (typically uncritically-held) beliefs and assumptions about the qualities and character of the relevant group. Nevertheless, there are contexts in which it is considered permissible to make judgements about

individuals as members of independently identifiable social groups, on the basis of beliefs and assumptions about the nature of those groups, which lead to their members being systematically disadvantaged in some way, which we do not consider discriminatory. Insurers, for example, make economic judgements about individuals, as members of independently identifiable social groups; and when they do this, they act in a manner that often seems to invite charges of unfairness. At the same time, such judgements are not generally considered to be discriminatory. Thus the practices of the insurance industry might initially seem to be a fertile source of counterexamples to the interpretation of discrimination developed in the last chapter.

If my motor insurance company assigns me to a high-risk, high-premium category, it will do so by treating me as a member of at least one readily-identifiable social group. The interpretation of discrimination given in the previous section may seem to suggest that, in any case in which such an assessment looks unfair, I should, given the nature of the relevant decision-making process, be understood to be a victim of discrimination. But while it is certainly true that insurers often arrive at such decisions in a manner that invites charges of unfairness, we would typically hesitate to accuse them of discrimination. How then can discrimination be distinguished from simple unfairness in such cases?

Being judged discriminatory is in fact only one of the many ways in which an insurance assessment decision might be judged to be unacceptable. Perhaps the simplest type of unacceptable case is that in which the assignment of an individual to a particular risk category involves a straightforward error: the individual is assigned to (e.g.) a high-risk category, to which she does not in fact belong. A slightly more challenging case of miscategorisation would be one in which, rather than simply assigning an individual to the wrong group, the company mistook the nature of the group itself, correctly assigning an individual to a group that was falsely understood to be (e.g.) high-risk.

Both of the above cases involve straightforward errors, and for a company to charge a higher premium on the basis of an error of this sort would reasonably be regarded as unacceptable. However, these are by no means the only possible bases on which an insurer's decisions or policies might be questioned. A more interesting type of case is that in which the above errors are avoided, and the relevant generalisations concerning risk (that the individual concerned belongs to such and such a group, and that the group is typically high-risk) are correct as generalisations, but are false in the particular case. As an example of this type of case consider the situation of a competent and conscientious teenage driver, 'correctly' assigned to a high-risk category. In this sort of case we may well feel that the individual concerned is the victim of unfair treatment. But is she a victim of any sort of injustice? In the last chapter the presence of some form of procedural unfairness was considered to be the hallmark of an unjust distributive decision. On this basis however, it is hard to view those correctly assigned to high-risk groups, even where they are not personally high-risk, as the victims of any kind of injustice. The insurance company is obliged to operate in a procedurally fair manner – the assessment process may not be loaded, nor rigged, against any individual or group. Thus the company is amongst other things under an obligation to treat like cases alike – it may not make

exceptions for individuals. In the company's defence then, it could be argued that it acts in the only way that is compatible with procedural fairness when, in the case in question, it charges our teenage driver a similar premium to others falling within the relevant age bracket.[1]

It might still be argued, of course, that the outcome in this case involves a distributive unfairness. But the insurance company could respond that this is one of those unfortunate cases in which, as also occurs in a properly run lottery, procedural fairness results in distributive unfairness. (They could also point out that the apparent unfairness in this case is only a particularly striking example of the more widespread unfairness to which moderately unsafe teenage drivers are subjected when they are lumped together for assessment purposes with the downright dangerous – a pattern that is as a matter of fact duplicated across every other age range.) Where the company has fulfilled its duty to act in a procedurally fair manner then, the outcome of such an assessment looks not to involve any form of injustice, and a fortiori not to involve discrimination, even in cases in which it appears to involve significant unfairness.

It is indeed hard to see the victims of the above sort of unfairness as victims of discrimination except in the following type of case: suppose I am correctly identified as a member of a genuinely high-risk group, but the causes of the group's being high-risk are known to lie in certain highly regrettable features of the wider social context – for instance in a range of unjust and oppressive social institutions and practices, which have operated over a considerable historical period. In this type of case, while my membership of the group in question may be incontestable, as may the fact that the group is in general high-risk, there will be a recognisable element of general social responsibility for the group's high-risk nature. In such cases we may well see individuals who are members of a group that is already widely understood to be disadvantaged, standing to be further disadvantaged on the basis of an arguably unfair assessment process, which forms the basis for a distributive decision. In this sort of case an insurer who charged me a higher premium, on the basis of my membership of such a group, could reasonably be accused of discrimination. The source of concern in cases of this type, which leads us to suspect a form of more or less calculating procedural unfairness, is the fact that the relevant generalisations concerning the group reflect a history of disadvantage. Our underlying conviction in such cases is that people should not be assessed on the basis of their membership of such groups, when significant distributive decisions are being made – membership of such a group ought not to figure in the rationale for a distributive decision.

Why, though, should a history of disadvantage be so important – transforming what might otherwise be an acceptable form of commercial prudence into an unacceptable form of procedural unfairness, and social injustice? In some cases the relevant history of oppression/disadvantage may be associated with propaganda

---

[1]Of course, everything would then turn on what should be considered to be a 'like' case. But that is, in a way, my point. It is certainly possible for the insurer to argue that the relevant form of likeness is age, rather than driving ability.

and prejudice, the existence of which casts serious doubt on the reliability of any beliefs or assumptions concerning the group in question, which might be used as the basis for an assessment decision. If we have good reason to suspect that the relevant beliefs and assumptions may not be true, we have good reason to think that it would be wrong to base any distributive judgement upon them. Perhaps surprisingly however, the procedural unfairness involved in such an assessment is not purely, or even primarily, a function of the likely *unreliability* of any relevant beliefs and assumptions, in any case in which there is this sort of historical background. Members of such groups may be exposed to possible discrimination through being the victims of popular but false preconceptions. But they are not necessarily protected from discrimination in cases where the relevant preconceptions are actually true. Every generalisation or cluster of generalisations (such as that such and such a group is high-risk) presupposes a categorisation.[2] And while some such categorisation-generalisation complexes are considered suspect because they do not reliably reflect the facts about the group in question, others are properly considered suspect because, however accurately or inaccurately they may reflect the observable qualities and character of the group in question, we cannot rule out a general social responsibility for the relevant observable features of the group.

It would for example be considered discriminatory to charge members of particular ethnic or religious groups higher insurance premiums, even if research showed that the members of those groups were in fact relatively high-risk. This is presumably because, given a historical background of ethnic and religious strife and oppression, we cannot eliminate the possibility that we share a general social responsibility for any relevant factor, including a propensity to risk-taking, that might turn out to be a feature of the members of such groups. It is not that true and informative generalisations concerning (e.g.) risk behaviour could not in principle be made when classifying by ethnic origin and religious identification then; nor is it the case that such generalisations could not be meaningfully applied to individual cases. Rather, due to the highly complex and controversial character of the social connections between these factors and risk behaviour, and the background of a history of ethnic injustice and religious intolerance, it simply is not considered fair or appropriate in most cases to categorise by (e.g.) ethnic background or religious affiliation when assessing likely risk behaviour.[3]

In such cases we have to say that bringing certain categorisation-generalisation complexes into play is for many purposes simply illegitimate, however true and

---

[2]Henceforth I will use the phrase 'categorisation-generalisation complex' to refer to the clusters of beliefs and assumptions which, allied to a particular social 'taxonomy', form the basis for decisions of this kind – which are also in play in, for example, social scientific studies of the causes of injustice.

[3]That said, it may be acceptable to bring ethnic origin into play in relation to e.g. health insurance, where there is a firmly established and fundamentally asocial connection between ethnicity and certain health conditions. But this tends to reinforce the general point. Only where the connection between ethnicity and risk factors is asocial do we feel that it is not discriminatory to take it into account.

informative the relevant generalisations may be. That is: however relevant a particular categorisation-generalisation complex may seem to be on a narrow view, many simply *should* not be made the basis of individual assessments, since to use them as the basis for such assessments would be to disadvantage the members of such groups on the basis of factors for which society in general is thought to bear significant responsibility. Any procedure that crucially depends upon such a categorisation-generalisation complex will count as procedurally unfair whether or not any other form of rigging is evident. And where there is this sort of background of socio-historical distortion, it would be discriminatory to allow the relevant categorisation-generalisation complex to play a role in the assessment decision, since to do so would be to subject the individual concerned to a more or less calculating procedural unfairness that targets her 'as' a member of an independently identifiable group.[4]

Where the legitimacy of the categorisation-generalisation complex used as a basis for assessment is the important factor, when determining whether a given case constitutes a case of rigging, it is noteworthy that the question whether the individual involved does in fact exhibit the (e.g.) high-risk character of the group in terms of which she is assessed seems relatively unimportant. Earlier, in considering the justice of the case of the safe teenage driver, the fact that the individual concerned did not share the high-risk character typical of the group as a whole was the basis for the judgement that her treatment was unfair, but not unjust. However, once we see that, as discussed above, procedural fairness more or less guarantees that such cases will crop up (since it requires us to treat 'like' cases alike, in circumstances in which no two cases are in fact exactly alike), the fact that the individual concerned is actually a safe driver looks less significant. Given that the insurer must assess individuals as members of *some* group(s), cases like this will inevitably arise. And if we have no reason to think that the relevant individuals are the victims of procedural unfairness, we seem obliged to conclude that their treatment, though possibly unfair, is not unjust, and a fortiori not discriminatory.[5] On the other hand, if it would be procedurally unfair to assess me on the basis of my membership of a group whose high-risk nature is essentially a legacy of a history of oppression and strife, then – surprisingly perhaps – it remains discriminatory whether or not I am personally high-risk. The significant differences between these two types of case do not therefore revolve around the truth of the generalisations on the basis of which the risk level of the individual is assessed. In the genuinely discriminatory case, to assess

---

[4]With factors such as gender, ethnicity or sexual orientation we will equally tend to regard differential premiums as discriminatory unless the relatively high- (or low-) risk behaviour can be tied to some largely asocial causal mechanism. Care needs to be taken here however – I am certainly not *defending* the idea that while assessing by ethnicity would be discriminatory, assessing by (e.g.) gender is not.

[5]I am assuming for the sake of the example that younger drivers are typically high-risk, and that the causal basis for this high-risk nature is reasonably well understood, and predominantly asocial – having to do with inexperience, excitability etc., factors which are likely to be associated with youth under any social conditions.

the individual in question as a member of the relevant group would be discriminatory, whether or not she is, as an individual, high-risk. In the non-discriminatory case, the assessment is just even in a case in which the individual concerned does not share the high-risk nature typical of the group.

The above examples help illustrate the importance of taking account of what I have referred to as the 'legitimacy' of a categorisation-generalisation complex, in cases of possible discrimination. But what makes for the legitimacy of such a complex? Is it perhaps only complexes that reflect a history of oppression etc. that are properly considered illegitimate? That these are not the only possible illegitimate complexes is evident when we consider examples such as the following.

I happen to belong to the group of people whose first names and family names have the same number of letters. Now, it is conceivable that accident figures might show a strong statistical correlation between people with this property and high-risk behaviour.[6] Thus a situation could conceivably arise in which an insurer wished to charge me a high premium on the basis of my membership of this group. But this would surely be unjust. Even where there is a statistical correlation between involvement in accidents and having the same number of letters in one's first name and family name; and even supposing that I am, as an individual, unusually likely to be involved in an accident in the near future; the relative number of letters in my first and last names could not legitimately be used as the basis for a fair assessment of my risk status. This is because, in the absence of any evidence of a causal link between length of name and high-risk behaviour, the likelihood of my being involved in an accident could not, even in the face of considerable statistical evidence, be reliably established by reference to this factor. Procedural fairness in making the relevant risk assessments thus requires more than just ensuring that the group-membership of individuals is accurately established, and that the relevant generalisations relating to the risk profile of different groups are true. It also requires that the categorisation-generalisation complexes utilised are both legitimate in the broader historical sense we have already considered, and are such as to yield both predictive and explanatory power. If this is not the case then procedural unfairness will result. And where this leads to the systematic and reasonably foreseeable disadvantaging of an independently identifiable group (as opposed to a mere assortment of individuals – as in the length-of-names case), the result will be not simply injustice but discrimination.

4.2 Injustice and Stereotyping

We have found then that in any case in which individuals are assessed as members of groups, as the basis of the rationale for some distributive decision, we can fully guard against discrimination only by ensuring that the categorisation-generalisation

---

[6]The statistical correlation could be a matter of sheer coincidence, or it could reflect the presence of some genuinely relevant factor that a subgroup of members of this group share with each other, and with others, who don't belong to the group at all.

complexes used in the assessment are legitimate. On the one hand the grouping must be such as to pick out genuine groups of individuals, rather than more or less random assortments, whose only significant common features are either irrelevant, or are directly connected with the assessment criteria. On the other hand, we must be able to show that the relevant features of the groups in question have a reasonably secure, relatively asocial, causal basis, and are not simply socio-historical constructs, reflecting a regrettable history of oppression and strife. That is, we must be confident that the categorisation-generalisation complex captures something pertaining to the underlying *nature* of the group in question, prior to any significant social distortion. If the relevant categorisation-generalisation complex simply records what a group has become, as a result of processes and events that should never have happened, it would be procedurally unfair to make use of it, even where it yields considerable predictive and explanatory power.

The error made by those who fail to ensure procedural fairness in cases such as the length-of-names insurance example discussed above could be described as a fallacy of relevance. There are a vast number of potential ways to categorise human beings, and it may turn out that some of these are correlated in surprising ways with socially or economically significant trends. But when judging others, especially in ways that affect their fundamental rights and well-being, and economic fortunes, we typically insist that they be judged only the basis of categorisations that are legitimate, and relevant to the matter in hand. Importantly, it is not primarily a question of the existence of evidence of an observable trend linking the group in question to some socially or economically significant factor. The important question is whether the classification picks up on some independently relevant feature that serves to (partly) capture the nature of the group in question. Justice in such cases demands not simply abstract fairness, but that in the administration of a fair procedure, we judge individuals according to categories that are explanatorily and predictively relevant, and are non-arbitrary. Of course existing anti-discrimination legislation is in practice able to prevent many possible forms of discrimination, in the insurance industry and elsewhere. The usefulness of the insurance example however lies in the way that it reveals that (1) not all such unfairness qualifies as discrimination, and (2) discrimination need not involve any straightforward falsehood or miscategorisation, but may involve the illegitimate utilisation of true generalisations in a manner that systematically disadvantages members of particular groups, and adds to and compounds more general disadvantages they already suffer.

If one seeks to capture the difference between discrimination and other forms of injustice stemming from procedural unfairness, the important factor seems to be that discrimination rests in part on (conscious or unconscious) *stereotyping*. A discriminatory practice doesn't simply treat people in a manner that is procedurally unfair; it treats them unfairly on the basis of certain stereotypical notions of who and what they are. Thus the concepts of discrimination and stereotyping are closely linked, and in cases where discrimination is suspected, concerns about stereotyping will also be pertinent.

The term 'stereotype' originates as a name for the blocks of solid type etc. formerly used in printing. A stereotyped judgement is thus, by an easily comprehensible metaphor, a judgement about others that is delivered in a pre-formed and automatic manner, with no appreciable sensitivity to individual cases. Concern about stereotyping may in some cases focus on the fact that the generalisations on which the stereotyped judgement is based are not actually true – it simply is not the case that most or all of the members of the relevant group share the particular (typically undesirable) trait that the stereotyped judgement implies. In this case we might speak of a false stereotype – the stereotype presents a false picture of the group in question. Alternatively, the concern may focus not on the correctness of the stereotype as a generalisation, but on the fact that there will almost always be exceptions to the general rule. Even if the generalisations a stereotype embodies *are* true of the relevant group as generalisations, the stereotype may always turn out to be false in the particular case before us.

Each of the above concerns proved relevant to the insurance case. Where the generalisation on which the assessment was based proved to be false as a generalisation, we said that the individual concerned seemed to have strong grounds on which to protest his or her treatment. The second type of concern arose in the case of the teenage driver. Here though, while the driver concerned would seem to have some grounds to claim that the result of her assessment was unfair, we would tend to defend such stereotyping in practice, as an unavoidable consequence of the insurance company's attempt to fulfil its obligations in respect of procedural fairness. In some cases then, the observation that the stereotyped judgement is untrue will serve as an adequate basis on which to contest the judgement. In others, perhaps surprisingly, it will prove inadequate. The value of considering the latter type of case – in which a generalisation proves to be untrue in the particular case – is that it draws our attention to the question of how we are to determine what counts as a 'like' case. It may seem obvious that treating all members of a group equally (or 'weighting' them equally in instances where the relevant assessment is multi-factorial, as it will typically be in practice), is an example of 'treating like cases alike'. But as our earlier discussion has shown, unless the use of the relevant categorisation-generalisation complex is itself legitimate in the context in question, the basis of any supposed 'likeness' will be questionable. While the teenage driver cannot plausibly claim, simply on the basis of the falsity of the relevant generalisations in her particular case, that she is a victim of discrimination, she may still be entitled to object on the basis that the age-based classification used lacks all legitimacy. (She might for example reasonably protest that she should not be assessed in terms of her likeness to teenage drivers in general, but in terms of her likeness to members of some other independently identifiable group of typically safe drivers, whether teenagers or not.) On this basis she could claim that she is being stereotyped – not because teenagers are safer drivers than is usually assumed, nor even because what is true of teenage drivers in general is untrue in her particular case, but simply because categorising her by age, rather than by some other category, is wrong-headed. This is a different sort of concern about stereotyping. It is not a concern about the truth of a stereotype, but about the legitimacy of the employment of the relevant categorisation-generalisation complex in the first place.

   As a further illustration of the point here consider again the case of people whose first and last names have the same number of letters. Suppose my insurer were to assess me as high-risk on the basis that research has shown that membership of this group is statistically correlated with high-risk behaviour. In such circumstances it seems beside the point to protest that the generalisation is untrue, either at the level of a generalisation, or in my particular case. The real problem is that the duty to treat like cases alike can only reasonably be understood to apply where the relevant 'likeness' provides a plausible basis for predicting and explaining my behaviour. If, as seems overwhelmingly probable, having an equal number of letters in one's first and last names has nothing whatever to do with a genuine propensity to engage in high-risk behaviour, it would be illegitimate to assess me under this heading, even in the face of a clear statistical correlation between the two factors. In the absence of evidence of some reasonably comprehensible mechanism linking relative length of names to high-risk behaviour I can protest that the statistic alone provides no basis for a reliable assessment in my case, and I am being judged according to a meaningless stereotype. The duty to treat like cases alike only makes sense then against the background assumption that stereotyping is to be avoided – that the business of deciding what counts as a like case will be carried out responsibly. Where the initial categorisation is botched, any related cluster of generalisations will, even if true, only add up (supposing they have any longevity at all), to so many stereotypes.

4.3 Ideological Stereotyping

Notwithstanding the above points, we noted earlier that some categorisation-generalisation complexes *can* yield predictive and explanatory power, and still fail to qualify as legitimate bases on which to assess people. In cases of this type, it was the fact that the underlying mechanisms involved reflected a high degree of social arbitrariness (e.g. through a history of oppression), which undermined their legitimacy. In these cases too – indeed especially in these cases – the victims would seem to have strong grounds to complain about stereotyping. Thus the full picture is as follows: complaints about stereotyping may raise concerns about the truth of generalisations, pitched either at the general level, or at the level of the applicability of the true generalisation in some particular case. At a deeper level however, they may also raise concerns about the legitimacy of a given categorisation-generalisation complex. This may be because the relevant categorisation-generalisation complex lacks predictive and explanatory power. Or it may be because, while the categorisation-generalisation complex does possess predictive and explanatory power, its doing so is simply a consequence of a regrettable history of social injustice, oppression etc., and tells us little or nothing about the true nature of members of the group in question.

   Granting that concerns about 'stereotyping' can properly arise from issues relating to the (claimed) untruth of a generalisation, at either the general or the individual level, it is possible to distinguish what I shall call 'ideological stereotyping' from

stereotyping more generally, using this term to refer to those cases that raise concerns about the legitimacy of a categorisation-generalisation complex (based either on suspicions of predictive and explanatory irrelevance, or on suspicions of historical distortion). It is this type of stereotyping that is key to understanding discrimination. In genuine cases of discrimination a distributive (or other) decision is made, the procedural unfairness of which rests on, or is (consciously or unconsciously) reinforced by, an ideological stereotype.

Although basing an unfair distribution of social benefits or harms on generalisations that are untrue at either the general or the particular level is, prima facie, a source of legitimate concern, it is important to recognise that we are often prepared to tolerate decisions arrived at in this way, which in some cases will be the result of adhering to relevant forms of procedural fairness (as in the case of the safe teenage driver). By contrast, it is often hard for people to appreciate why we should not assess individuals on the basis of certain stereotypical notions, where the generalisations on which the stereotypes are based seem to be true, or at least reasonably reliable. Nevertheless, as the preceding discussion has aimed to show, there are good reasons to 'bracket' or 'put out of play' certain stereotypical notions in certain circumstances, because of concerns about their *ultimate* predictive and explanatory power. Some stereotypes may fail to yield any explanatory or predictive power at all. Others may yield predictive power under prevailing socio-historical conditions, but fail to be genuinely explanatory, because they do not reflect the true nature and potentialities of the group in question. Arising from oppressive and distorting historical circumstances, they record only what a particular group has historically been permitted to become.

Ideological stereotypes are a fertile basis for offences against persons. Earlier I presented discrimination as involving a combination of injustice and 'rigging', where rigging of the relevant (group-oriented) variety involves an ideological element. We are now in a position to refine this conclusion somewhat. The victim of discrimination suffers an additional indignity, by comparison with the victim of simple injustice. Discrimination arises when injustice is combined with procedural or structural unfairness on the basis of an ideological stereotype. In discrimination, the victim of injustice suffers additionally. But this is not because she suffers an additional 'harm', which can somehow be added to the harm she suffers though being the victim of (e.g.) distributive injustice. Rather, it makes more sense to say that to be the victim of stereotyping, where this involves being the victim of discrimination, is in part at least to suffer a wrong, which cannot be straightforwardly aggregated to the harm due to the associated distributive injustice. It is conceivable that one might stereotype someone while acting beneficently toward them – as is no doubt the case in many instances of medical paternalism. To act paternalistically is typically to wrong an individual while pursuing her interests, and through the manner in which one pursues her interests. But it hardly makes sense to say that paternalism harms while it benefits, for this fails to capture the sense in which in acting paternalistically one may genuinely pursue an individual's interests in an unmixed and unambiguous fashion, and yet still wrong her through the manner in which this is done.

   To be the victim of stereotyping has more in common with being the victim of a lack of respect for one's autonomy than it does with being the victim of either maleficence or distributive injustice. It is one's human dignity that is compromised, rather than necessarily one's 'well-being', narrowly conceived. (Though of course in cases of discrimination in which stereotyping is combined with *distributive* injustice one may be both wronged and harmed). It follows that, as in a case of failure to respect autonomy, one can 'suffer' as the victim of stereotyping – which is essentially an affront to one's human dignity – without necessarily being harmed in the process.

   It might be thought that no new ethical approach is needed to address the problem of ideological stereotyping, either on its own, or – as discrimination – in conjunction with rigging-type injustice. But these concerns about stereotyping, and the associated threat to human dignity, highlight how inadequate contemporary theoretical conceptions of 'respect for persons' are, in this respect. To stereotype someone is to fail to respect her personhood in a sense that is not reducible to a failure to respect her choices, or a failure of distributive fairness. Everyday ethical and political discourse implicitly recognises that we have responsibilities to persons – such as the responsibility to avoid stereotyping them – that are not accounted for on the basis of conventional notions of justice or respect for autonomy. Stereotyping is not reducible to a failure to respect the capacity for self-determination, conventionally understood – or for that matter to any form of straightforward injustice – because it concerns in part the basis of, and the kinds of admissible evidence for, judgements concerning injustice. One can behave, in a narrow sense, 'respectfully', or beneficently, toward someone while still being guilty of stereotyping her, as is the case in much paternalistic behaviour. Stereotyping, like discrimination, to which it is often closely linked, is an interpretive moral wrong. One does not simply stereotype someone, one stereotypes her 'as' a member of this or that group. It is however important not to think of ideological stereotyping as necessarily occurring as either an element in discrimination, or in some 'free-floating' form. Stereotyping can also be of ethical significance via its connection with a different but related moral wrong: instrumentalism. As we will see in the next chapter, when stereotyping is combined with instrumentalism the result is *objectification*.

Chapter 5
Objectification

Abstract This chapter argues that objectification constitutes a third interpretive moral wrong, alongside discrimination and stereotyping. Objectification combines stereotyping with instrumentalisation. Given the fragmentation of the moral person, discussed earlier, the view that the instrumentalisation of ontological persons is uniquely unacceptable might seem to rest on questionable grounds. Nevertheless, in view of the self-interpreting capacities of ontological persons, there is a distinctive wrong associated with reducing them to the status of a means: that of inducing them to adopt an instrumentalised self-conception. The wrong of objectification consists not simply in what we do to another, but in how we interpret them, and how we thereby invite them to think of themselves. There is potential for a form of reduction to the status of a means here which goes beyond anything that need result simply from treating another as a means. An examination of the phenomenon of sexual objectification serves to highlight some of the distinctive features of this form of reduction to the status of a means.

5.1 First-Stage Objectification: Instrumentalisation

As noted in the Introduction, 'objectification' does not currently seem to be a very fashionable ethical concept. It also does not seem to be a very well-understood concept; and perhaps these two facts are connected. In this and the following chapter I set out a three-stage analysis of the concept, which is designed to bring out both its complexity, and its broad ethical significance.

One area in which the idea of objectification is still very much alive is in the ethics of biomedical research. Here though, the term tends to be used in a relatively restricted sense – a diminution of a much richer ethical idea. Typically, efforts to ensure that the subjects of scientific research are not 'objectified' are geared to ensuring that due respect is shown for the 'autonomy' ('individual' autonomy, in O'Neill's sense) of research subjects.[1] The principle of respect for autonomy is held to have special importance in the context of biomedical research because, by contrast with the clinical context, to which general principles of biomedical ethics have

---

[1] See O'Neill 2002.

tended to be tailored, it is often (arguably, *always*) the case that the subjects of biomedical research have no personal interest in serving as research subjects.[2]

In the clinical context it is understood that the requirements of the principle of respect for autonomy must be balanced against the requirements of other fundamental principles, such as the principle of beneficence. This is because, given certain possible configurations of the patient's wishes, the patient's state of mind, and his or her basic interests, it may be appropriate in some cases to override the requirements of the principle of respect for autonomy in order to fulfil the requirements of the principle of beneficence (or others).[3] In the case of research however, things stand differently. In cases in which it is not in an individual's interests to serve as a research subject (as in a great many cases it will not be) it is simple exploitation to involve them in research without their consent (or, where appropriate, some plausible substitute for consent). Orthodox interpretations of the principle of beneficence in medical ethics treat the interests of the individual patient as paramount.[4] And where an individual has no interest in serving as a research subject, there can be no question of overriding her refusal to participate by appeal to the principle of beneficence, since in any such case there can be no argument from beneficence for the inclusion of that particular individual in the research study at all.

Moreover, even in cases in which the subject does stand to benefit from the research, it may well not be in her interests to be included in the study. In cases in which an individual suffers from the condition that the study aims to investigate, it may well be in her interests that the research is carried out. But it does not follow that it will be in her interests to participate in it. From any point of view that puts the interests of the individual patient first, it will make more sense for her to allow others to run any associated risks, so that she may then reap any resulting benefits.

The principle of respect for autonomy thus plays a very different role in the research context from that which it plays in the treatment context. Instead of functioning primarily as a check on paternalism, its role is to provide researchers with the only satisfactory warrant they can have for including a given subject in research in the first place. Only the subject's free choice to participate can entitle us, in the face of the burdens and risks associated with participation in research, to proceed to involve her in a course of action which may be at best tangential to her own interests.

Granting for the sake of argument that obtaining adequate consent is sufficient to guarantee that the subjects of research are safe from inappropriate exploitation then (though there are doubts about this – see below), why might it be thought that consent is a reliable safeguard against *objectification*? The answer is that in the medical ethics literature, 'objectification' tends to be interpreted as a synonym for *instrumentalisation*.

We have already considered the Kantian idea that showing proper respect for persons involves refusing to reduce them to the status of a mere means. To reduce

---

[2]See Lucas 2010.

[3]Beauchamp and Childress 2001, p. 104.

[4]Ibid., pp. 165–76.

a person to the status of a mere means is to treat them simply as an object of use. In what is perhaps its most familiar use, the term 'objectification' refers to this sort of instrumentalisation of persons. However, as we shall see, this is by no means the only, or indeed the most illuminating, use of the term.

While it may be morally wrong, from a Kantian perspective, to reduce a fellow human person to the status of a mere means, it is an unavoidable fact of life that we must frequently employ others as means in some capacity or other. Whenever we pay someone to work for us, or we work for another, one person serves as a means for another. However, it is arguable that, in ordinary circumstances, we are not reduced to the status of a mere means when we engage in waged labour. In fact, the distinction between slavery and waged labour can shed light on the difference between serving as a means, and being reduced to the status of a mere means. Waged labourers are arguably not reduced to the status of a mere means when they are employed on a contractual basis, are rewarded for their labour, and are free to dispose of their earnings outside of working hours, as they see fit. Slavery, by contrast, is the classic example of the instrumentalisation (that is, 'objectification', in the sense currently under discussion) of another human being. Slaves are not party to any contract, and they earn nothing through their labour. They are regarded as property, to be bought, utilised and disposed of as the owner wishes.

As a first approximation, the contrast between slavery and waged labour does seem helpful as a way to shed light on the distinction between using others as means, and reducing them to the status of mere means. However, closer inspection suggests that the distinction between slavery and waged labour is not as clear-cut as we might imagine. Although in many cases those engaged in waged labour are able to preserve an adequate measure of autonomy, in others it is doubtful whether this can be done. In well-regulated industries in broadly egalitarian contexts it is no doubt possible to draw a sufficiently clear line between using an employee as a means, and reducing them to the status of a mere means. But this is certainly not true in all cases. Many examples could be given of cases in which waged labourers cannot really be said to have freely entered into their contracts of employment, or be free to dispose of their earnings as they choose outside their work hours. A slave owner may consider the slave's welfare only insofar as it affects his own interests, but often an employer's concern for his employees does not look so very different from this. In such cases the status of the employee may not be far removed from that of a slave – making expressions such as 'wage slavery' seem far from oxymoronic. Such cases illustrate how questionable the supposed distinction between serving as a means and serving as a mere means may be in practice. Prostitution provides another example of where the line between serving as a means and serving as a mere means becomes hard to draw. Even in the least coercive contexts, it seems evident that the fact that the prostitute enters into a notional contract, and receives payment, is not really sufficient to exonerate the client from the charge of reducing a fellow human to the status of a mere means.

Nevertheless, it is possible to defend the distinction between using someone as a means, and reducing them to the status of a mere means, by focusing on the question whether the employment in question satisfies the goals of the employee, in addition

to those of the employer. If my labour serves as a means to my employer's goals, and simultaneously serves as a means to my own goals, then I can plausibly be said to serve as a means, while not being reduced to the status of a mere means. On the other hand, if the nature of my employment is such that none of my own goals is satisfied through it, we seem justified in speaking of my reduction to the status of a mere means.

The above way of shoring up the means/mere means distinction, by focusing on the question whether the relevant employment also satisfies the employee's own goals, has been utilised in the context of avoiding objectification in biomedical research. [5] The existing guidance and literature on the ethics of biomedical research tends to focus on the importance of adhering to the principle of respect for autonomy, as a means by which the inappropriately instrumental treatment of research subjects can be avoided. Of course, even the fully informed and voluntary consent of the research subjects cannot change the fact that they are, in a quite straightforward manner, a means to the researcher's ends. Nevertheless, the distinction between the use of another as a means, and reducing them to the status of a mere means, can still mark an ethically significant boundary. Thus Leonardo de Castro argues that where the goals of the research subjects coincide with those of the researcher, or more particularly where the research subjects have identified with the goals of the research, and adopted those goals as their own, no question of inappropriate instrumentalisation – 'objectification', in the sense currently under discussion – arises. Only where participation fulfils none of the research subjects' own goals will they be reduced to the status of a mere means. [6]

5.2   Second-Stage Objectification: Adoption of Alien Goals

We have seen that instrumentalisation is avoidable where the goals of the employee are satisfied via employment (through being paid, for example, or where they or coincide with, or come to coincide with, the goals of the employer). However, on this matter of the potential coincidence of the goals of the employee with those of the employer, it is necessary to note a couple of important qualifications.

Firstly, it is necessary that the goals adopted by the employee are such that they could rationally be adopted by a human being. Again, biomedical research provides an illustration of the potential problems that can arise here. While the research subjects' endorsement of the goals of the research is surely a necessary condition for their involvement in research to be ethically acceptable, it is doubtful whether – even combined with consent – such an endorsement is sufficient for the research to be ethical. The concern with the possible instrumentalisation of research subjects is founded on a concern to protect the subject's *dignity*. It would be absurd to suggest that serving as a research subject is in general incompatible with preserving one's

---

[5]See de Castro 1998.

[6]Ibid., p. 381.

human dignity. The volunteer who understands the goals of the research, endorses those goals, and generously (in some cases perhaps even heroically) agrees to participate, in the face of significant risks, is not thereby stripped of human dignity. On the contrary, we might well regard this sort of altruism as exemplifying the human capacity to act in a freely self-determining manner, rather than being a slave to self-interest and pleasure seeking. However, it is equally evident that there are some possible uses of human research subjects that would be incompatible with protecting their dignity, even where the researcher had obtained the subject's informed and fully voluntary consent. In at least one notorious case, in which the informed and voluntary consent of the research subject was given, there was widespread public revulsion at the idea of reducing a human being to the status of a vivisectional animal, simply in order to hasten the progress of biomedical research.[7] Evidently, we cannot rationally consent to just any type of treatment and preserve our dignity intact – and in principle this applies equally to cases in which the subjects have adopted the researcher's goals as their own. Thus what matters is not simply what research subjects are, as a matter of fact, disposed to consent to, but whether the goals they share with the researcher are such that they could rationally be adopted by any research subject.

The second important qualification also relates to the nature of the goals adopted by the employee; and this qualification has the capacity to shed significant further light on the phenomenon of objectification. Thus far we have treated the question of whether an individual serves as a mere means as if what were in question were only the explicit terms on which that individual is required to serve. We said that if the employee's role is such that only the goals of the employer are satisfied through it, the employee is thereby reduced to the status of a mere means. By contrast, if at least some of the employee's own goals are thereby satisfied, they may serve as a means, but (we said), they do not serve as a mere means.[8]

The above distinction remains workable enough in contexts where there is no question of interplay between the goals of the employer and the goals of the employee. If we assume that each of them approaches the situation with settled goals, and with those goals clearly in view, we can, on the basis of an examination of whose goals will be realised in and through the relevant process, come to an accurate view of whether the employee is being required to serve as a mere means. But of course the goals of human beings are rarely fully settled or fully conscious. We do not travel through life with certain fixed goals constantly in view. Although many of the goals associated with our continued biological existence can reasonably be taken to be broadly settled, the widespread human tendency to physiological self-harm indicates that very often we do not even have those goals very clearly in view. Moreover, human beings have a characteristic propensity for embracing new goals, and rejecting old ones. They may even invent entirely new goals for themselves (as, for example, the goals associated with mountain climbing were invented in the fairly

---

[7]The case of John F. Russell, in New Jersey in 1902. See Lederer 1995, pp. 24–5.

[8]Though it should be noted that the prostitution case represents a possible counter-example.

recent historical past). Thus it is unrealistic to assume that there can be no interplay between the goals of the employer and the goals of the employee.

This plasticity of human goals permits an employee to serve, on occasion, in a role that might otherwise have been straightforwardly exploitative, without risk of reduction to the status of a mere means. If I adopt some of the goals of my employer, I may thereby avoid reduction to the status of a mere means. My employment now serves my own goals, just as it serves those of my employer, not simply because I receive financial or other recompense for my trouble, but because some of the goals of my employer have now become my own. But this business of the adoption of goals by the employee can cut both ways. I might equally well reject some of my pre-existing goals when I begin to appreciate how, given prevailing economic arrangements and opportunities, they serve to enmesh me with the activities of an unscrupulous employer. (For example, if I want to start a family, but find that the only way to provide for a family financially would be to accept employment on highly exploitative terms, I might choose to abandon my goal of having a family.) Indeed, the fact that goals can be embraced and rejected in this manner is central to the idea that the instrumentalisation of research subjects can be avoided through their coming to identify with the goals of the researcher.[9]

It cannot reasonably be denied then that we must allow for interplay between the goals of the employer and those of the employee. But with this in mind, the idea that instrumentalisation is avoided where the employee adopts the employer's goals as her own begins to look naïve. It is a commonplace that an employee may come to identify with the goals of her employer to a degree that is ultimately disadvantageous to her. Moreover it is evident that when this occurs the employee is not only likely to suffer some degree of harm, but, more particularly, is likely to be *increasingly* reduced to the status of an instrument. A contract that was 'freely' entered into is no proof against one's reduction to the status of a mere means in a case in which one has become so heavily identified with the goals of one's employer that one has lost sight of one's more personal guiding values. As has become apparent in many cases of industrial upheaval, the fact that a worker identifies strongly with the goals of the employer is no protection against redundancy – with the worker being 'let go' like an obsolete piece of machinery. What we see in such cases is the substantial instrumentalisation of the employee, notwithstanding any freely-entered-into employment contract. In such circumstances, the employee's identification with the goals of the employer does not serve to stave off a reduction to the status of a means. On the contrary, it tends to reveal just how thoroughly instrumentalised the employee has become. It is through adopting the goals of the employer that the employee becomes reduced to the status of a cog in a potentially obsolete machine.

A broadly similar argument can be made in relation to prostitution. It may seem natural to contrast the instrumentalisation of a prostitute with the autonomous

---

[9]A key aspect of de Castro's discussion, for example, is that the goals of the research subject are to some extent plastic: the subject is assumed to be capable of adopting certain of the researcher's goals, and the ethical acceptability of the research is held to hinge partly on this possibility.

conduct of those involved in consensual sex. But on the one hand the prostitute's position has meaningful parallels with waged labour (where it doesn't amount to sex slavery – though as has already been argued the boundaries here are not clear cut). While on the other hand consensual sexual activity cannot be guaranteed to be free of a kind of false consciousness similar to that which we find in the case of the employee who over-identifies with the goals of the employer. As many feminist thinkers have emphasised, it is not safe to assume that women who enter into 'consensual' sexual relationships with men always do so freely. Far from being evidence of emancipation, the fact that a woman has adopted the goals and mores of a male-dominated society to the extent of 'freely' participating in 'normal' sexual activity may actually be a reliable index of her oppression.

We may conclude then that the fact that an individual has adopted the goals of another is not sufficient to show that instrumentalisation is absent. On the contrary, it may be evidence of just how complete that individual's instrumentalisation ('objectification', in this sense) has become. This suggests that the instrumentalisation of as complex a creature as a human being is never a simple affair. If we were the sorts of beings that came equipped with a restricted set of instinctual goals, and if these goals were evident to any reasonably attentive and unbiased observer, then it might be a simple matter to determine which of a given range of employments amounted to the reduction of the employee to the status of a mere means. But that is not how human beings are. And the very factors that make the adoption of hitherto alien goals possible for us simultaneously serve to make any actual case of adoption ambiguous. Epistemically, we may reasonably ask: how can we be sure that an employee has in fact adopted the goals they or others claim they have? Ontologically, we may reasonably ask: do the goals that have been adopted belong authentically to that employee as their own? This ontological concern takes two forms: Were the goals in question adopted in the right sort of way, such that they constitute firm and abiding commitments on the employee's part, rather than fleeting whims? And, supposing the answer to this latter question is yes, are the goals in question compatible with the most basic underlying goals of the individual concerned, such that they belong to them authentically, and are not a manifestation of some sort of syndrome of self-harm and/or self-betrayal? Only if we could say with confidence that the goals adopted belong to the research subject authentically could we reasonably fall in with de Castro's suggestion for avoiding instrumentalisation. We will return to this question of the authenticity of the employee's adoption of the relevant goals in due course. Before doing so, it is necessary to complete our survey of the stages of objectification by considering more explicitly the role played by reflection and self-interpretation in the reduction of an individual to the status of a means.

5.3   Third-Stage Objectification: 'Reduction' and Reflection

At this point it is helpful to return to Downie and Telfer's discussion of the principle of respect for persons, to consider a point they make concerning the ways that the development of a rational will can be undermined or restricted by the actions

of others. Downie and Telfer note that not only our capacity for action, but also our *conception* of our capacity for action, can be adversely affected by the actions of others. The development and form of the personality has, they observe, a great deal to do with how one is able to conceive of oneself. It is a commonplace that we may undermine or restrict another's capacity to act when, through an overly solicitous attitude to them, we refuse to allow them to make the sorts of mistakes that are essential to certain kinds of learning.[10] But it is also important, though less frequently noted, that we may undermine or restrict personal development by failing to show due respect for the other's capacity to form an adequate self-conception:

> The development of personality can also be blocked on a grander scale by political arrangements which restrict the range of images which people can form of themselves.[11]

Any restriction on the range of images people can form of themselves can have consequences for their potential objectification. In relation to first-stage objectification – objectification as simple instrumentalisation – the image that the 'master' forms of the 'slave' may serve to legitimate the master's conduct, at least from his own point of view. Constructing a particular image of an oppressed group is one way to legitimate, amongst the oppressors, the continued oppression of that group: 'these people are really nothing but resources, put here for our use, and in using them as instruments we only do with them that which they are most naturally suited to'.

However, as the above quotation from Downie and Telfer indicates, there may be a further consequence of such a restriction, above and beyond the legitimation of instrumentalisation among the instrumentalising group. In second-stage objectification, discussed above, the images of the instrumentalised group promulgated by the instrumentalising group may influence the instrumentalised, in a manner that facilitates their ongoing instrumentalisation. By encouraging the victim to adopt alien goals (for example, by encouraging the slave to adopt the goals of the master), such images bolster and perpetuate the instrumentalisation of the victim. This role played by 'images' in second-stage objectification requires us to take great care when considering the supposed autonomy of the victim. In a case in which the victim identifies heavily with the image promulgated by the instrumentalising group, it is likely that any supposedly autonomous choices will not be genuinely autonomous at all (since they will not be informed by the victim's authentic goals). In this respect then, a capacity to endorse the images promulgated by an oppressor may restrict the full development of the personality, by restricting the development of a capacity for free autonomous activity. Our earlier example of the identification of the employee with the goals of the employer shows how this might occur. The victim of second-stage objectification isn't simply reduced to an object of use, but becomes an *educable* object of use, who can be brought to adopt certain alien goals. The possibility of reducing a person to an object in this sense resides precisely in

---

[10]Downie and Telfer 1969, pp. 20–1.

[11]Ibid., p. 21

the fact she is *not* simply a thing, but has certain qualities that are distinctive of persons. She is not merely an instinctive being, bound to act in pursuit of various biologically-conditioned goals, but is the sort of being who can adopt new goals, and develop new strategies for their realisation. If, then, she can be brought to adopt the goals of the oppressor, the relevant higher-level cognitive capacities can be recruited to her own detriment. She can be brought to act 'freely' and choose 'autonomously' in conformity with the oppressor's instrumentalising agenda.

It is this latter possibility that seems to be the prime focus of Downie and Telfer's observation that the development of personality can be restricted by controlling the images that people are able to form of themselves: by manipulating those images one can produce in any being with a sufficiently developed personality a propensity to act in pursuit of alien goals, in a manner that in turn restricts the *full* development of its personality. This emphasis on the potential consequences for the restriction of people's capacity to act with genuine autonomy, by restricting the self-images available to them, is in line with the general orientation of Downie and Telfer's analysis, and indeed with the traditional moral theories on which they draw, which tend to construe autonomy, and associated notions of the healthy development of the personality, as if the central concern for any ethic of respect for persons should be our potential freedom of *action*.[12]

To take our analysis of objectification further however, it is important to appreciate that an oppressor's control of the images his victims are able to form of themselves may have ethical ramifications above and beyond its potential effects on the victim's freedom of action. Objectification need not simply affect the victim's capacity to act in pursuit of her own goals. It may also affect her self-conception directly. The development of her personality would thereby be blocked in a highly distinctive way. The adoption of an instrumentalised and alien self-conception would affect not only the victim's capacity to act autonomously, but also, more fundamentally, her capacity for undistorted self-interpretation – thus for self-knowledge. In this latter case, in addition to suffering through being conscripted to serve as an instrument in the furtherance of another's goals, she would suffer through being brought to *assent to* a particular instrumentally derived conception of what she is. This would be an ethically significant consequence both because it would facilitate the victim's ongoing instrumentalisation, and because the free exercise of this capacity is, arguably, an important element in our ontological personhood in and of itself.

This further dimension of objectification is hinted at, but not explicitly developed, by Downie and Telfer, when they refer to the blocking of the development of personality through restricting the types of images that people can form of themselves. We can regard it as a further stage of objectification – *third-stage* objectification. It would have its own distinctive ethical significance insofar as our free development

---

[12]This focus upon autonomous action is for example distinctive of Kant's moral philosophy: '*Autonomy* is therefore the ground of the dignity of human nature and of every rational nature' (Kant 1996, p. 85).

as human beings involves not only the development and exercise of our capacity
to act in pursuit of goals that are genuinely our own (autonomy, in the traditional
sense), but also the free development and exercise of our capacity for undistorted
self-understanding – self-knowledge. Later I will outline a case for the claim that
respecting the free development and exercise of the latter capacity is central to
respecting human dignity – and thus to any adequate version of a principle of respect
for persons. Before doing so however, I want to consider an example of third-stage
objectification, drawn from the work of feminist theorist Andrea Dworkin.

5.4 Andrea Dworkin on Sexual Objectification

As indicated above, third-stage objectification relates not so much to our capacity
for goal-setting and pursuit, as to our capacity for undistorted self-interpretation –
though of course these capacities may be inseparable in practice, insofar as any
process of goal-setting must involve an element of conscious self-interpretation.
This dimension of objectification is particularly closely connected with the 'problem
of objectification' considered earlier (in connection with Lorraine Code's discussion
of individuals whose personhood is in doubt – see Section 2.3.3). However, it can be
more effectively illustrated with reference to Andrea Dworkin's analysis of sexual
objectification.

   Andrea Dworkin was, and understood herself to be, a *radical* feminist thinker.
Her definition of a political radical, following Octavio Paz and Ortega y Gassett,
is that of someone who is not simply concerned to correct contemporary abuses,
but in addition seeks to critique and correct contemporary 'uses'.[13] On this view,
the radical's task is that of getting us to see that the normal is not normative – that
there is, in at least some cases, a hidden continuum between what we find abnormal
and objectionable, and what we find normal and acceptable (and this with a view to
problematising the normal, rather than rehabilitating the abnormal).

   In *Pornography: Men Possessing Women*, Dworkin's focus is the hidden contin-
uum between sexual objectification and sexual fetishism.[14] It is well known that
in some people sexual responses become tied to various kinds of 'inappropriate'
objects. Men (Dworkin concentrates on heterosexual male sexual fetishism, and
henceforth I will do the same) may become fascinated with some particular part
of women's bodies – for example feet, or hair. Or they may become fascinated with
some accessory or item of clothing – shoes or underwear. The fascination becomes
a fetishisation when the object comes to provoke a fixed sexual response.[15] The
fetishising male may become sexually aroused simply by being in the presence of
the fetish object. For example, the mere sight of a woman's bare feet, or underwear,
comes to be exciting in and of itself, even in contexts devoid of any other sexual

---

[13]Dworkin 1988, p. 147.

[14]See Dworkin 1981.

[15]Ibid., pp. 123–4.

dimension or reference. In more developed forms the fetish object may cease to be merely a sufficient condition of sexual arousal, and become a necessary condition as well: the man is unable to sustain his sexual arousal except in the presence of the relevant object.[16]

A consequence of the development of a fetishised sexual response, Dworkin observes, is that the man's sexual behaviour becomes increasingly divorced from any tie to particular sexual partners. Whether or not a given sexual relationship is a stable one, or a long-term one, it will be undermined to the extent that the male partner finds himself reacting with sexual arousal to any suitable object – to any naked foot, or item of underwear, for example. And of course in the more developed form, in which the presence of the object becomes a necessary condition of sexual arousal, any sexual partner is likely to become painfully aware that the man is not responding sexually to *her* at all, but to the fetish object. At this extreme the partner will be aware that she has effectively been supplanted. Although she may remain involved in her partner's sexual activity, she cannot any longer pretend to be central to it. It is the fetish object that is now the primary focus. The individual flesh and blood partner may remain a desirable feature in some respects, but she is henceforth ultimately a dispensable one.[17]

In some of its forms fetishistic behaviour may be trivial, and almost comical. In others it may be extremely sinister and disturbing. But in any case, there is, in Dworkin's view, something pathological about it.[18] It is, according to Dworkin, impossible to regard fetishistic behaviour, especially in its extreme forms, as a healthy or desirable feature of an individual's sexual life. To the extent that we regard a degree of stability as a desirable thing in sexual relationships, it is hard not to sympathise to some degree with Dworkin's view. Fetishistic behaviour seems inevitably to compromise the stability of sexual relationships, since it shifts the focus from the partner to the fetish object.

Dworkin highlights the constricting character of sexual fetishism when she considers the ways in which it may be necessary for the sexual partner to play a particular role, even perform to what is in effect a prepared script. These cases reveal something essential to all sexual fetishism, which is that it is *scripted* behaviour.[19] Apologists may wish to stress and celebrate the peculiar capacity of human beings to invest ordinary objects with extraordinary power; and this sort of presentation of fetishism may suggest that the fetishist inhabits a world of extraordinary sexual possibilities. But this appearance is, in Dworkin's view, deceptive. In fact, the scripted nature of the fetishist's sexual responses betokens anything but an opening up of sexual possibilities. Rather, fetishism bears all the marks of compulsive, stereotyped

---

[16]See in this connection D.H. Lawrence's comments on the young men of his day: 'To them, sex means just plainly and simply, a lady's underclothing, and the fumbling therewith' (Lawrence 1971, p. 338).

[17]Dworkin 1981, chapter 4.

[18]Ibid.

[19]Ibid., pp. 110–1, 124–7.

behaviour. The fetishist who is fixated with women's shoes for example, cannot help but miss out on many of the possibilities afforded by fully human relationships with the people who wear them. Or consider Gustave Flaubert's apparently fetishistic response to prostitutes. What may appear, from some perspectives, to be an uninhibited and life-affirming response to the sexual possibilities of ordinary lives looks on closer inspection to be highly constricting. A man who has an automatic sexual response to the mere sight of a prostitute is at constant risk of distraction from the more important relationships in his life. To the extent that he automatically responds to someone whose sexual services are simply for sale, his partner might reasonably doubt whether the sexual dimension of their relationship has any real importance for him. And in the more extreme case of the fetishist who *only* responds sexually to prostitutes, the constricting character of the fetish is vividly apparent. It may be that the fetishist can travel the world paying for the sex he wants with any number of partners. But if sex is only possible for him in the context of a financial transaction, the wider possibilities of meaningful human sexual relationships are clearly closed off.[20]

Thus, for Dworkin, fetishistic behaviour vastly constricts ordinary human sexual relationships. While from the fetishist's standpoint there may not appear to be any constriction ('surely sex is not about love, intimacy, shared passion, trust, fidelity, it is about *shoes*!'), if we view the situation from a more detached standpoint there is something sad and stunted about even the most harmless forms of fetishism, to say nothing of the more sinister and disturbing forms. Dworkin's next step is to argue that there is no clear line between sexual fetishism and sexual objectification. Dworkin initially glosses 'objectification' in accordance with its use in sexual psychology, where it refers to, in effect, the male fetishisation of another human being: 'that fixed response to the form of another that has as its inevitable consequence erection'.[21] However, she goes on to note that the interpretation of objectification offered by sexual psychologists involves a simplification and a narrowing-down of a term with much broader connotations. The phenomenon of objectification, on Dworkin's account, has three prominent features. Firstly, men objectify women when they treat them as mere instruments for their own use and pleasure.[22] Secondly, the way in which women are so treated involves an element of fetishisation. The woman is treated in such a way as to provoke, in the fetishist, an automatic, and from his point of view an unproblematic response.[23] This response is tied to its object in such a way that it ensures that she cannot refuse to provoke it. For this reason the process of objectification bypasses her subjectivity – she has no effective choice in the matter. Thirdly, and apparently connected to the element of fetishisation, the real focus of objectifying behaviour is not the victim herself, but

---

[20] Ibid., pp. 127–8.

[21] Ibid., p. 113.

[22] Ibid., pp. 108–10.

[23] Ibid., pp. 111–5.

an *ideal* that she serves to symbolise. For an illustration of this idealising element in objectification, Dworkin draws on Thomas Hardy's *The Well-Beloved*:

> To his Well-Beloved he had always been faithful, but she had had many embodiments. Each individuality known as Lucy, Jane, Flora, Evangeline or whatnot had been merely a transient condition of her. . . Essentially she was perhaps of no tangible substance; a spirit, a dream, a frenzy, a conception, an aroma, an epitomised sex, a light of the eye, a parting of the lips.[24]

To the objectifying male, what matters is not the individual woman, but something that transcends her, in the name of which she might at any moment be discarded.

Thus characterised, sexual objectification amounts to the sexual stereotyping and fetishistic instrumentalisation of women. Dworkin regards sexual objectification of women, thus understood, as a pervasive feature of our culture. Women are encouraged to conform to sexual stereotypes from an early age, and through a variety of pressures. Clothing, make-up, behaviour – how women move, speak, work and amuse themselves – very largely involve conformity to a male-authored script. In some cases the connection is obvious. Some female roles revolve around male conceptions of attractiveness, and pander to them. But even some archetypes of apparent liberation – the go-getting career woman, in charge of her own life, with her own disposable income, and a series of disposable sexual partners – may figure in the aspirations of teenage girls not as a straightforward role model, but as a model of female attractiveness *to men*. The teenager wants to conform to the archetype not because it is attractive to her in itself, but because it epitomises what she takes female attractiveness to be about.

For most ordinary women, most of the time, according to Dworkin, life is a struggle to become a particular type of object: the object of male sexual desire, which has certain hallmarks, and behaves in certain stereotyped ways.[25] Just as male-to-female transsexuals go to great lengths to 'pass' as women, so too *women* spend much of their time trying to pass as women – struggling to play the part assigned them by men, on which much of their financial security, and personal satisfaction, have come to depend.

Dworkin regards pornography as the primary vehicle of this objectifying process.[26] Although objectifying mechanisms pervade our culture in general, from the fashion industry to the news media, from pop music to advertising, the themes they reproduce are held to originate in pornography. It is in pornography that the overtly sexual basis of the relevant stereotypes is most apparent. It is also in pornography, especially but not exclusively in its less respectable forms, that the fact of instrumentalisation is most apparent. In these cases women are not reduced to the status of objects simply in virtue of being coerced to play along with a script written by men, for the satisfaction of men. Rather, they are literally reduced to the status of objects. Dehumanised, used and abused as objects – their status as thinking, feeling

---

[24]Thomas Hardy, quoted in Dworkin 1981, p. 114.

[25]Dworkin 1981, pp. 124–8.

[26]Ibid., p. 128.

human beings is explicitly denied.[27] According to Dworkin, pornography reveals the truth that underlies all sexual objectification – all forms of male scripting of women's behaviour. Its basis is fundamentally (though perversely) sexual, and its purpose is the fixed (fetishised) male sexual response to the woman-as-object. This is what links the most violent and repulsive forms of contemporary pornography to literary 'classics', such as the work of the Marquis de Sade, to the 'soft' pornography of men's magazines, to pop videos, advertising, and fashion magazines and comics aimed at teenage girls and pre-teens.[28]

Not only is male sexual objectification of women pervasive then, it is, according to Dworkin, indistinguishable in its form from sexual fetishism. Referring back to her conception of the political radical, her approach is a radical one because she aims to reveal the pervasive male objectification of women as one element in a hidden continuum that also incorporates more widely recognised forms of fetishism. Sexual objectification and sexual fetishism are not separated as 'use' and 'abuse' – as they might appear to be to liberal commentators, concerned only to remedy apparent abuses. Not only is there a hidden continuum linking pornography in all its forms to advertising and pop videos, but this continuum is supported by a virtually universal male fetishisation. Men may think they are fascinated by women. In fact the object of their fascination is a particular sexual object, whose role they have scripted, and as which most women (and some men) spend much of their lives trying to pass. For Dworkin, to the extent that the presence of this 'object' is a necessary condition of male sexual arousal, all men display fetishistic behaviour, of a fairly developed and extreme kind.

In her campaigning work against pornography Dworkin chose to focus less on these theoretical complexities than on the actual physical and emotional damage that is done to women directly in the production, distribution and consumption of pornographic material.[29] Dworkin's message in *Pornography* is a slightly different one: in contemporary society virtually all men are sexual fetishists; as a consequence healthy sexual relationships are practically impossible, and human possibilities are hugely constricted – seriously damaging many lives, male and female alike. Some forms of fetishism are obvious enough, but the obviousness of these forms of fetishism serves to cloak the most widespread and damaging form: the virtually universal male fetishisation of women themselves as nothing more than sexual objects.

The idea that almost all of us conspire in the sexual objectification of women, may seem overstated. On the account given thus far it may well seem that the 'us' here must be understood to apply to men, or possibly even a subset of men, and that it would be better restricted to this usage. On Dworkin's more developed view however, this is not necessarily the case. In her later book *Intercourse* (1988) she

---

[27] Ibid., chapter 6 and passim.

[28] Ibid., pp. 66–9.

[29] A version of Dworkin's definition of pornography (developed in conjunction with Catherine McKinnon) was incorporated into Canadian obscenity law in 1992.

emphasises that the sexual fetishisation of women amounts to something more than a conspiracy on the part of men. Women display a certain complicity in their own objectification.

In *Intercourse* Dworkin characterises sexual objectification in now-familiar terms as a precondition for male sexual performance.[30] She emphasises that there is a need for the female partner to conform to certain behaviours and scripts preordained by men. However, in her view there is nothing straightforwardly biological about this need, nor about women's widespread conformity with these behaviours and scripts. Dworkin insists that humans are not instinct driven, to this extent. Rather, we inhabit a world of ideas and meanings. We make choices within this world of meanings, and the meanings available to us actively delimit the choices we make. We might flatter ourselves that we choose freely, but the truth is that the widespread male sexual objectification of women creates the vocabulary within which men and women alike make their choices.[31] Thus the pervasive male sexual objectification of women prevents any individual act of sexual 'surrender' being an example of a free choice.[32]

However, Dworkin argues, the initial surrender – the surrender to the system of meanings that objectification resides in and depends upon – *does* involve an element of freedom. Once the surrender is made, it is nonsense to suggest that women freely consent to sex. But the initial surrender – the choice to become the object that the system of male sexual objectification demands – is not inescapable. This is where Dworkin sees real potential for change, since, in principle at least, women can refuse to play along with their objectification. But alongside this element of freedom comes a corresponding responsibility. By playing along with male sexual objectification women are not just subject to domination, they are implicated in the domination they suffer. Women, Dworkin claims, 'collaborate' in their own objectification to the extent that they fall in with, and fail to challenge, the objectification system that requires their submission.[33] Thus, Dworkin concludes, it is appropriate to view the objectification system as something in which virtually all of us – female as well as male – conspire. And the very universality of the conspiracy simultaneously makes it difficult to detect. Male domination of women looks much of the time to be something that women themselves demand and perpetuate, and (consequently), not like domination at all. The brilliance of objectification as a strategy, Dworkin observes, is that it gets women to take the initiative in their own oppression.[34] However, it remains domination because although it involves acts of individual self-policing, it is built on a female self-conception that is a construction of the dominant sex. The proximal vehicle of domination may in many cases be a woman's own beliefs and

---

[30] Dworkin 1988, p. 148.

[31] Ibid., pp. 164–5.

[32] Ibid., p. 165.

[33] Ibid., p. 167.

[34] Ibid., p. 169.

attitudes. But the source of domination is the male demand that she conform to the characteristics of a particular type of sexual object.

It is here that the importance of the 'idealising' element in objectification – which goes beyond merely treating another as an object – becomes apparent. What is it about the particular form of 'looking beyond' the physical body to an intangible spirit – exemplified by Hardy's 'Well-Beloved' – that is so important? Although this element in objectification is not inconsistent with the fetishisation of women that is the primary focus of Dworkin's attention in *Pornography* – the fetish is an object certainly, but a magical object, with mysterious powers, not just a brute thing – Dworkin does not there fully explain the significance of what her discussion hints at. However in *Intercourse* the shift of emphasis from what men do and have done to women, to women's reception of and reaction to what is done to them, enables Dworkin to explain how women come to collaborate in the process. On the one hand, male sexual objectification of women remains a fetishisation. Women are not reduced to mere objects, but are treated as special, almost magical objects. The woman-as-object is not a mere thing, 'it' points beyond itself, and refers to a host of associated objects, practices, behaviours, desires. On the other hand, of course, the women who are the objects of this fetishisation are in reality much more than objects. They have minds as well as bodies, they have the capacity for self-understanding, and the exercise of this capacity *contributes to* their objectification.

No doubt there is, from the point of view of the objectifying male, a special frisson associated with fetishising a person, rather than a mere thing. But a further and most important consequence is the fact that, as a self-knowing 'object', the victim is capable of collaborating in her own objectification. On Dworkin's fully-developed view, women adopt a self-conception that is modelled on a male-authored stereotype, and adapt their behaviour to suit. Thus the 'idealising' element in objectification is not simply a decorative flourish by which the objectifying male further entertains himself. It is a key element by means of which the most complete form of objectification is accomplished. Men turn women into fetish objects, and they do so in order to procure an unproblematic sexual response in themselves. However, this is done in the full knowledge that women are much more than objects – they are subjects who (some of the time at least) very much want to be capable of provoking that response. So, men treat as fetish objects non-objects who partly wish to be fetish objects. But the price of becoming, temporarily, and on strictly controlled terms, this magical object, is the victim's complicity in the 'truth' that *this is what she is*. The reduction to the status of an object does not come about through the *denial* of the woman's subjectivity, but through the *occupation* and surrender of her subjectivity. Her victory is her defeat, the price of the male acknowledgement of her as a sexually attractive being is her acquiescence to the 'truth' that she is essentially a sex object.[35] She isn't allowed simply to play the role of a thing, she must also assent to the implicit claim that she is a thing. And the spice of the procedure, from

---

[35]Ibid., pp. 166–7.

the point of view of the objectifying male, is that, knowing that she knows that she is in reality not merely this sex object, he nonetheless extorts from her an implicit acceptance that this is what she is.

Objectification thus understood is coercive, but its coerciveness is relatively subtle. Rather than simply circumscribing its victims' scope for free movement, as it undoubtedly does to some degree, male sexual objectification, on Dworkin's account, affects women at the level of their sense of identity. Even in the absence of overt coercion, it has a coercive effect, by scripting women's behaviour, and moulding their sense of who and what they are, in such a way that even apparently free choices are constrained. Indeed, women are trapped in a double bind – apparent fulfilment lies in self-abasement: I will respect you only on the condition that you surrender your self-respect, leaving me the option of refusing to show you any respect, as it suits me, and justifying this by pointing to your own evident lack of self-respect.

5.5 Third-Stage Objectification as an Interpretive Moral Wrong

Dworkin's radical claims raise a number of critical questions. I will not attempt to address such questions in detail now, though I will return to some of them in the next chapter, in which a critical discussion of Dworkin's claims will serve as a springboard to another phase of our enquiry. For the present, I wish to round off and summarise this chapter's findings, with Dworkin's example of sexual objectification in mind.

I have in this chapter distinguished three different 'stages' of objectification. First-stage objectification is simple instrumentalisation. Second-stage objectification involves, in addition to instrumentalisation, the adoption by the instrumentalised party of the instrumentaliser's alien goals. Third-stage objectification – of which sexual objectification on Dworkin's account would be an instance – involves a distinctive element: the moulding and colonisation of the victim's sense of her own identity. In third-stage objectification the instrumentalised party is not simply treated as a tool, or as a resource. She is treated as a tool or resource in a manner that also involves her adoption of a distorted, stereotyped and instrumentalising self-conception.

Referring back to our earlier discussion of interpretive moral wrongs, it is clear that first-stage objectification is not an interpretive moral wrong. One may be treated as a tool or a resource without being so treated under some description – without being instrumentally stereotyped 'as' a member of this or that group.

Equally clearly, third-stage objectification, as characterised above, is an interpretive moral wrong. One cannot be objectified in a manner that involves the adoption of a stereotyped and instrumentalising self-conception except by being brought to do so *under some description*. Third-stage objectification therefore ranks alongside discrimination and (simple) stereotyping as a moral wrong in which the victim

is wronged 'as' a member of this or that group. For example, one could only be sexually objectified, in Dworkin's sense, as a woman.

How then should we understand second-stage objectification – is it also an interpretive moral wrong? Since it is conceivable that someone might be brought to adopt alien goals by a process that did not involve the adoption of a stereotyped and instrumentalising self-conception, it seems necessary to acknowledge that there may be some forms of second-stage objectification that do not qualify as interpretive moral wrongs. Nevertheless, it seems likely that in many cases the most effective process by which an individual could be brought to adopt such goals would involve the adoption of such a self-conception. Whether this means we should say that there are also instances of second-stage objectification that do qualify as interpretive moral wrongs, or whether it means we should say that there may be many instances of what might initially appear to be second-stage objectification that would be better regarded as instances of third-stage objectification, is a question that we might be permitted, for the moment, to leave open. Having considered various ways in which the reduction of a person to the status of means is in part accomplished through their adoption of a stereotyped and instrumentalising self-conception, it does not seem necessary to preserve a particularly hard and fast distinction between our second and third stages. What matters more is that we remain alert to the importance of such stereotyping where it is in evidence. It is in these cases that we would be justified in speaking of objectification as an interpretive moral wrong.

References

Beauchamp, Tom L., and James F. Childress. 2001. *Principles of biomedical ethics*, 5th edition. Oxford: Oxford University Press.

de Castro, Leonardo. 1998. Ethical issues in human experimentation. In *A companion to bioethics*, eds. Helga Kuhse and Peter Singer, 379–389. Oxford: Blackwell.

Downie, R.S., and Elizabeth Telfer. 1969. *Respect for persons.* London: Allen and Unwin.

Dworkin, Andrea. 1981. *Pornography: Men possessing women.* London: Women's Press.

Dworkin, Andrea. 1988. *Intercourse.* London: Arrow Books.

Kant, Immanuel. 1996. Groundwork of the metaphysics of morals. In *Immanuel Kant: practical philosophy*, ed. Mary Gregor, 41–108. Cambridge: Cambridge University Press.

Lawrence, D.H. 1971. A propos of Lady Chatterley's lover. In *D.H. Lawrence: A selection from Phoenix*, ed. A.A.H. Inglis, 327–361. London: Penguin.

Lederer, Susan. 1995. *Subjected to science.* Baltimore, MD: John Hopkins University Press.

Lucas, Peter. 2010. Is 'Therapeutic research' a misnomer? In *Arguments and analysis in bioethics*, ed. M. Häyry et al., 229–239. Amsterdam: Rodopi.

O'Neill, Onora. 2002. *Autonomy and trust in bioethics.* Cambridge: Cambridge University Press.

Chapter 6
Interpretive Moral Wrongs and Radical Theorising

Abstract The account of objectification offered in Chapter 5 raises a number of questions concerning the relationship between the analysis of interpretive moral wrongs and radical theorising. This chapter aims to show, via a range of examples (Martha Nussbaum's reformist attempt to analyse sexual objectification; Marx's analysis of commodification; and the ethics of genetic databases), that an ethic of self-interpretation, suited to tackle interpretive moral wrongs, must adopt a radical rather than a reformist theoretical approach. The chapter, and Part I of the book, closes by drawing together themes from the preceding four chapters, to arrive at an interim conclusion, concerning the normative grounds of an ethic of self-interpretation. Kant held that respecting the dignity of human persons requires us to respect above all their capacity for principled self-determining action. I argue that respect for ontological persons also crucially involves respecting their dignity, through respecting their capacity for undistorted self-interpretation, and self-knowledge. This points the way to a distinctive and enduring role for an ethical principle of respect for ontological persons.

6.1 Dworkin's Radicalism

The previous chapter identified a form of objectification – third-stage objectification – in which the instrumentalisation of an individual or group is accomplished through their adoption of a stereotyped self-conception. In third-stage objectification individuals are reduced to the status of a means, not simply through being treated as an instrument or a resource, but through being brought to understanding themselves as instruments or resources.

Sexual objectification, as analysed by Andrea Dworkin, was given as an illustration of objectification, thus understood. However, I also noted that Dworkin's analysis gives rise to a number of critical questions. These questions concern the defensibility of some of her central claims, and in this chapter a consideration of such questions will help advance our understanding of the relationship between the analysis of interpretive moral wrongs and *radical* theorising.

Earlier chapters dealt with the interpretive moral wrongs of discrimination and stereotyping. Though the nature of these wrongs may not be perfectly understood, their existence is widely recognised. Many countries around the world have established legal safeguards to prevent discrimination on grounds of gender, ethnicity etc.

These help give us a working understanding of what discrimination is, and how it can be avoided. They also project a reasonably clear picture of what a society lacking in discrimination would be like. Broadly similar points could be made about stereotyping. Unlike discrimination, stereotyping is not the main focus of specific legal provisions. Nevertheless, in the context of professional ethics, the idea that we should be alert to the possible stereotyping of clients, students, patients and pupils etc. is a reasonably familiar one.

By comparison, the implications of Dworkin's analysis (and, potentially, the implications of any analysis of any variety of third-stage objectification) may well seem less clear. We might reasonably wonder how sexual objectification could ever effectively be challenged, if it is as pervasive as Dworkin claims. We might also wonder what a world without sexual objectification would be like. Before concluding that there is something inherently vague and ambiguous about third-stage objectification though – and in particular about sexual objectification as an example of third-stage objectification – it behoves us to consider whether it really is significantly more puzzling than other forms of objectification.

An orthodox Kantian ethic of respect for persons provides us with a formula for avoiding the reduction of a person to the status of a means (first-stage objectification). We are to act only on the basis of maxims that can rationally be universalised. It also provides us with a sketch of a society in which personhood is accorded due respect, in the shape of Kant's kingdom of ends (see Section 2.2). It might be tempting then to conclude that, unlike third-stage objectification, first-stage objectification is quite well understood. But is this really the case? The early sections of the previous chapter were geared to showing that once we take into account the plasticity of human goals, and associated difficulties in distinguishing between using another as a means and reducing them to the status of a mere means, it is much more difficult than we might have thought to say how pervasive first-stage objectification really is. It also turns out to be hard to say with any specificity what a society without such objectification would be like. Can we, for example, be sure which forms of waged labour (if any) would be acceptable in the Kantian kingdom of ends? Or can we be sure what forms (if any) of pornography, or advertising, it would permit?

It might be said that the above line of argument implies less that we understand sexual objectification tolerably well, than that we understand the Kantian notion of the reduction of a person to the status of a means rather less well than we might have supposed. But this is in itself a noteworthy result. My argument in Chapter 3 proceeded by pointing up some difficulties and puzzles surrounding accepted ideas about discrimination. I hope that I have succeeded in showing how some of these puzzles might be rendered less puzzling. But it should not be forgotten that, at a societal level, the relative clarity we all now enjoy regarding discrimination has been hard won. There is abundant evidence that nineteenth-century commentators found the idea of sex discrimination quite as baffling as some contemporary commentators have found the concept of sexual objectification. Of course, none of this adds further detail or clarity to the concept of sexual objectification, beyond what may have been achieved in the previous chapter. But my aim at this point lies in another direction. What I now wish to emphasise is that, firstly, the apparent difficulties surrounding

Dworkin's concept of sexual objectification are less a reflection of any vagueness in her account than they are a reflection of certain fundamental features of her general approach; and, secondly, notwithstanding these difficulties, in the context of an enquiry into interpretive moral wrongs, her general approach has a good deal to recommend it.

One possible way to defend some of Dworkin's key claims would involve invoking a distinction between the empirical and the conceptual dimension of her analysis. Dworkin is evidently concerned to challenge common sense beliefs about e.g. sexual behaviour. For example, her claim that statistically normal male sexual responses involve a form of fetishism challenges the common sense belief that sexual fetishism, however mild, is statistically abnormal. At the same time, her approach is clearly calculated to raise conceptual questions. The questions of what we should understand by 'normal' behaviour, and of how what is statistically normal might relate to what is genuinely normative, are implicitly posed in her discussion. Accordingly, it might seem possible to bracket the empirical dimension of her analysis, pending further investigation, and concentrate instead on the conceptual dimension – asking, first and foremost, what light her analysis sheds on the *meaning* of 'objectification'. On the basis that empirical claims about (e.g.) sexually normal behaviour are in any case premised on certain implicit conceptual claims, we might even argue that little of real importance in Dworkin's analysis would be sacrificed by proceeding in this way. But there are independent reasons for thinking that this would not be an appropriate strategy.

While it is true that we cannot definitively settle any empirical question without some clear idea of how associated conceptual questions should be settled, the fact is that we do standardly take many empirical questions to be decidable without prior investigation of associated concepts. Moreover, while there is some value in highlighting the fact that an enquiry such as Dworkin's can advance our understanding of a concept like objectification independent of any associated empirical questions, we will want to be persuaded that it has at least some empirical promise before we attach any great importance to whatever conceptual implications it may turn out to have. The real worry though is that separating the empirical from the conceptual dimensions of such an analysis would ultimately appear facile.

In introducing Dworkin's account of sexual objectification I characterised her as a theoretical radical. The radical theorist, I said, sets out to challenge accepted views of the use/abuse distinction. Accordingly, she is not simply advancing a set of empirical claims. But neither is she simply advancing a set of conceptual claims. Nor, finally, is she simply doing both at once. To the extent that an analysis lends itself to an easy separation of its empirical and conceptual dimensions, we will tend to find that each dimension leads us back to common sense views. An easily-separable empirical dimension leads us back to common sense because, once conceptual problems have been hygienically set aside, any remaining questions will appear to be empirically resolvable – in principle at least. An easily-separable conceptual dimension leads us back to common sense because, as long as we treat conceptual questions as matters of sheer logic, or of established linguistic usage, otherwise divorced from any reference to the way things are, we will tend to fall

back on ordinary language, or some equivalent, in resolving them. But to ask genuinely radical questions about (e.g.) sexual normality, is to ask questions that do not permit us to refer without further ado to some established procedure of empirical testing (for what underwrites the test?), *or* to established linguistic usage (for what underwrites our current vocabulary?). The radical's task is necessarily complicated by the fact that she cannot appeal to common sense views, or to any existing consensus, in order to support the claim that such-and-such a practice, widely regarded as innocent, is in fact abusive. Nor can she appeal to any common sense view of how things might be done differently.

It is in my view easy to overstate the extent to which Dworkin leaves us in the dark on how sexual objectification might be avoided, or what a world without sexual objectification would be like. But to emphasise that Dworkin does give us reasonably clear hints on both of these points would be to risk obscuring the main point: what is disturbing to common sense in her analysis is not that she fails to provide evidence for her claims, or arguments for her conclusions – she does both of these things – but that her radical approach deliberately dismantles the common sense empirical and conceptual handholds we usually rely on in order to assess such claims. For example, her argument that statistically normal male behaviour is recognisably fetishistic, and, consequently (normatively) abnormal, derails any potential empirical or conceptual appeal to common sense views of sexual normality.

Dworkin's approach is thus opposed to any *reformist* approach to the analysis of objectification. While the reformist seeks to correct what common sense already recognises as abuses, the radical sets out to challenge common sense views of the use/abuse distinction. To appreciate the real value of Dworkin's approach, and in particular its promise for an enquiry into interpretive moral wrongs, it is helpful to contrast it with what can reasonably be characterised as a reformist approach to the analysis of sexual objectification.

6.1.1 Martha Nussbaum on Sexual Objectification

Martha Nussbaum is one of very few philosophers to have attempted to analyse the phenomenon of sexual objectification.[1] Nussbaum acknowledges that the concept of objectification is a rich and complex one. Indeed, she distinguishes no fewer than seven different notions that are involved in the idea of treating someone as an object.[2] However, this apparent complexity is lessened when we note that the majority of these notions of objectification consist of relatively subtle variations on the traditional theme of instrumentalisation (first-stage objectification). The issue of the subjectivity of the victim of objectification features in only one of Nussbaum's seven notions of objectification. This particular form of objectification consists in the active denial of another's subjectivity, it is:

---

[1]See Nussbaum 1995.

[2]Ibid., p. 257.

[T]he act of turning a creature whom in one dim corner of one's mind one knows to be human into a thing.[3]

In this sense, objectification is not the sheer instrumentalisation of another, with no regard for her subjectivity. On the contrary, some reference to the victim's subjectivity is essential to it, since it involves the *active* dehumanisation of the other. It is impossible to objectify someone in this sense without at least implicitly acknowledging that they are more than a mere thing. To this extent then Nussbaum goes beyond the excessively limiting view of objectification as simple instrumentalisation. At the same time, Nussbaum's portrayal of the way in which the subjectivity of the victim comes to figure in objectification remains curiously limited. She explicitly acknowledges Dworkin's contribution to the analysis of objectification. But what is for Dworkin the crux of the phenomenon – the reflective process by which a particular instrumentalised self-conception is taken up by the victim into her own self-understanding, thereby facilitating and apparently legitimating her instrumental treatment – simply does not feature in Nussbaum analysis. The victim's subjectivity is denied, but it is not – as it is for Dworkin – colonised. Nussbaum's objectifying agent may know in one dim corner of his mind that his victim is a human being, but the specifically human qualities of the victim, which equip her to make her own particular sense of the treatment to which she is subject, are not the focus of attention in anything like the way they are for Dworkin.

Aside from this single appearance of the idea of the denial of another's subjectivity, Nussbaum focuses exclusively on the use of another as a mere object. In this connection she finds, perhaps surprisingly, that some of the uses of another that can be ranked under the general heading of objectification are benign, even desirable. She draws on a range of popular and literary sources in support of the claim that treating others as objects can, in some of its forms at least, be a normal and healthy part of sexual activity. (Though it is worth noting that some of Nussbaum's examples suggest, rather oddly, that sexual intercourse *in itself* – that is, irrespective of any mutable features of the social/cultural context – amounts to the use of another as an object, which is a claim that Dworkin, for all her radicalism, never makes.)

Nussbaum's claim that some forms of objectification are desirable (while others are not) is symptomatic of her general approach, which seems geared as much to combating sexual puritanism as it is to combating exploitation and oppression. To this end, she seems concerned above all to develop an analysis of objectification that will enable us to recognise apparent abuses, while casting no undue suspicion on ordinary sexual behaviour.[4] In contrast to Dworkin's radical strategy then, Nussbaum's approach can reasonably be characterised as a reformist one – which acknowledges that sexual objectification can take abusive forms, while arguing that not all of the ways in which we might treat others as objects are ethically problematic. Simultaneously though, the rich and complex concept developed by Dworkin, and in particular its *radical* edge, slips into obscurity. The analysis of objectification

---

[3]Ibid., p. 281.
[4]Ibid., p. 257.

is levelled down into a familiar story concerning the rights and wrongs of instrumentalisation. The elements of reflection and self-interpretation that are central to Dworkin's account are studiously overlooked.

Unlike Dworkin, Nussbaum succeeds in locating sexual objectification in a familiar moral universe. This no doubt bolsters the initial plausibility of her account. But it does nothing to bring out the real depth and challenge of the phenomenon discussed by Dworkin. Moreover, it does not seem accidental that Nussbaum's reformist approach overlooks what for Dworkin is the central feature of sexual objectification: the impact on the victim's self-conception. If sexual objectification is a process that affects us (male and female alike) at the level of our self-understanding, a world without sexual objectification would not belong to any familiar moral universe. It is therefore important to avoid any temptation to try to ease the acceptance of Dworkin's analysis by presenting it as somehow less radical than its author intended it to be. That common sense finds it challenging is not of course a touchstone of its truth. But if common sense had found her analysis easy to digest this would certainly be evidence that it had fallen short of her radical aims.

Naturally, the above comparison of Nussbaum's analysis with that of Dworkin *is* geared, in a sense, to easing the acceptance of Dworkin's analysis – by suggesting that to analyse the phenomenon of sexual objectification adequately we will need to adopt something like the radical approach taken by Dworkin. I do not imagine that anyone who was not initially inclined to give any credence to Dworkin's claims will be swayed by this comparison alone. But they might conceivably be persuaded that there simply is no straightforward and commonsensical way to carry through her project. We should not expect any such analysis to deal as deftly and as plausibly with the phenomena it tackles as, say, an analogous analysis of discrimination or stereotyping might, since common sense views of both of these phenomena have been comprehensively softened up over the last century or so. However, the crucial point on which Nussbaum's reformist approach fell down was the acknowledgement of precisely that element in third-stage objectification that makes it an interpretive moral wrong; and there seems to me to be a general lesson here. Interpretive moral wrongs may, depending on the historical circumstances, demand radical analyses *because they engage us at the level of our self-conceptions*. To ask questions about them necessarily involves asking whether our familiar moral universe is really the place we have hitherto taken it to be.

By way of further illustration of the above points, I now wish to consider a further example of an interpretive moral wrong: the phenomenon of commodification, as analysed by Marx. As will become apparent, Dworkin's analysis of sexual objectification appears indebted to Marx's analysis of commodification on a number of points. Introducing the example of commodification at this juncture will therefore serve not only to underline the close connection I am claiming between the analysis of interpretive moral wrongs and radical theorising, it will also provide a further illustration of the sorts of difficulties theorists face when they attempt to provide neat and immediately plausible accounts of phenomena whose analysis requires that we challenge contemporary common sense.

6.2 Marx on Commodification

Under the economic conditions that prevailed in the cultures with which Marx was chiefly concerned, economic production was, paradigmatically, the production of commodities. By 'commodities' Marx meant goods considered as the objects of human wants. His particular concern however was that the value of commodities is typically analysed with reference to their exchange values – as commercial tokens, rather than as items of first-order use.[5] Classical economics analyses the patterns of human production, distribution, and consumption of commodities in a strictly 'scientific' manner. In doing so it identifies 'laws' of, for example, supply and demand. Such analyses may be presented as scientific; but in Marx's view they are scientific only in a restricted sense. They remain far too indebted to contemporary common sense, and as a consequence, the true nature of the commodity – its *essence* – is never investigated by classical economics.[6]

If the production, distribution and consumption of commodities had no implications for the well-being of the species as a whole, there might be little to be concerned about in this. But capitalist society's devotion to the production of commodities is, according to Marx, a perversion of human economic activity and social existence. The production of goods as commodities is an *alienated* and an *alienating* process.[7] The product that the worker produces is an alien and hostile power, whose sole purpose is to augment capital itself. And in its failure to carry through its analyses to grasp the nature of the commodity itself, classical economics is implicated in the debased and debasing historical conditions it describes. As the study of the laws of commodity production and distribution, which at the same time averts its gaze from the character of the commodity itself, classical economics is in Marx's view a science of alienated humanity – of humanity that has lost sight of what it is to be human.[8]

Since humans are naturally creative creatures, according to Marx, to initiate or help sustain a situation where the sole channel in which the creativity of large numbers of humans can express itself is the production of commodities is a form of cruelty. At one level the cruelty here is simply that involved in the systematic thwarting of a natural instinct – analogous to the way that a cow's instinct to suckle her young is thwarted by the extended and pointless lactation (pointless from the standpoint of the animals involved, at least) required by the modern dairy industry. Vast numbers of human beings are required to engage their naturally creative instincts and energies in the production of unnecessary and meaningless (to them) commodities. At another level however, commodity production involves a further distortion of peculiarly human capacities. Humans are not just creative, productive creatures.

---

[5]Marx 1954, volume 1, part 1, Chapter 1, section 1.

[6]Ibid., part 1, Chapter 1, section 4.

[7]Marx 1967, pp. 287–301.

[8]Ibid.

They are also naturally inclined to try to understand themselves and their world. And one important role for creative and productive activity is to aid in this process.

'Objectification' does not have the same sense for Marx as it has for Dworkin or Nussbaum. Instead, following Hegel, Marx uses the term to refer to the very general process by which human beings, through their creative or productive activities, reproduce themselves, and their consciousness, in things.[9] Thus craft activities, industrial processes, art, architecture – and so on across a whole range of human creative activities – centrally involve objectification. In this sense, an art object is an objectification of the artist. Of course the artwork could be a self-portrait, and in this case the idea of objectification has obvious application. The artist, we might quite naturally say, 'objectifies' herself in a self-portrait (and if the artwork is a three-dimensional sculpture the term seems particularly appropriate). But the important feature that makes the term 'objectification' appropriate is that, ideally, the artist should be in a position to recognise herself in the art object. The art object says something about the artist, and is both a reflection of, and a confirmation of, what she is. Of course it is not necessary that the self-portrait take a straightforwardly representational form for it to be appropriate to refer to it as an objectification of the artist. A more abstract work that reflects the artist's ethnic heritage, or religious commitments, may say as much about the artist as any representational piece. And just as a more or less abstract art object can count as a direct objectification of the artist, so a representational piece that is not intended to be a self-portrait can equally function as an objectification of the artist. A painting of the artist's children might say as much about the artist as a self-portrait would. A painting of the artist's patron could conceivably play a similar role, insofar as it records the world in which the artist lives, and does so through her eyes. Even a landscape might reflect something of the cultural and spiritual identity of the artist.

If even an abstract artwork or a landscape can be an objectification of the artist, insofar as it is, intentionally or unintentionally, a repository of personal and cultural attitudes and values, there is no reason to think that this phenomenon of objectification is confined only to artistic production. All the products of craft and manufacture to some degree mirror the circumstances, attitudes and values of the individuals and communities that produced them. The significance of objectification then, for Marx, resides in the fact that when we are engaged in objectification we are necessarily simultaneously engaged in implicit self-interpretation and self-description. The objects we create to some extent reflect back their makers. The architecture of a period, for example, can hardly fail to say something about the society that produced it. The same is true of the manufacture of tools, and of art and literature. Thus the history of human objectification can be read as a history of human self-interpretation – a history of human self-consciousness.

In artistic and economic production then, human beings are partly engaged in a process of self-interpretation. Thus although Marx's concept of objectification at first carries no negative connotations (indeed quite the reverse, a human being

---

[9]Ibid., pp. 331–3.

prevented from objectifying him/herself in some manner would be subject to a serious deprivation), in tandem with Marx's analysis of commodity production it becomes clear that objectification can take strongly negative forms. Human productive activity includes artistic production, and also the production of goods and the provision of services. In the production of commodities, where the significance of the product lies purely in its exchange value, the process of self-interpretation through productive activity is one that leads the worker to conceive of herself also as a commodity.[10] The commodity is not, like the best artworks, an object that reflects its producer, illuminating and confirming her sense of self. Instead it stands opposed to her, inscrutable and malevolent, an enemy and a cipher. In this case the worker's alienation from the object, reflected back, results in a self-alienation: we view our own productive and creative potential as a mere commodity, whose primary importance lies in its exchange value. When self-objectification through labour takes this commodified form, it results in a self-commodification – in Marx's view a profound misunderstanding of what it is to be human. And, of course, this self-misunderstanding is enshrined in classical economic theory, where human-labour-as-commodity is treated as if the ills associated with commodity production amounted to eternal truths about human nature.

It is a signal virtue of Marx's analysis of objectification and commodification that it recognises both of these levels of alienation. If, as a worker, I am simply 'commodified' by others, being valued for, and effectively reduced to, my labour power, I may suffer certain associated harms and wrongs. The process may involve both suffering and instrumentalisation – my reduction to the status of a means. But these results could conceivably follow without my coming to understand myself as a commodity – indeed they could occur without my having any very clear understanding of myself at all. However, human beings being the reflective and objectifying creatures they are, the historical transformation of economic activity into, primarily, the production of commodities, also begets the self-commodification of the worker. The classical economists' analysis of labour power on the model of the commodity plays a key role in the worker's coming to understand her own labour as a commodity. This outcome then has certain distinctive alienating effects. My intellectual development may be stunted, since I will tend to see no reason to develop my capacities beyond the level required for the most lucrative form of labour available to me. My creative faculties may also be stunted, not simply through the denial of opportunities, but also through my own conviction that what does not enhance my value in the labour market – whatever does not enhance the commodity I am – is of no value to me.[11]

Thus while Marx's concept of objectification differs significantly from those we have considered thus far, his conception of the process of the commodification of the worker, in which alienating forms of objectification are understood to play a key role provides a further illustration of the type of syndrome described by Dworkin (and hinted at by Code). When the worker's labour power is treated as a commodity

---

[10]Ibid., pp. 295–6.
[11]Ibid., pp. 290–3.

she suffers a profound instrumentalisation, which is simultaneously reflected in a stereotyped self-understanding. By this means an instrumentalising social arrangement is reinforced and legitimated through an ideological stereotype: a conception of labour as commodity that does not reflect the true nature of the worker, but rather what prevailing social conditions have made of her. When the stereotype, mediated by authoritative 'scientific' economic and social analyses, is taken up in the victim's self-understanding, it compounds the wrong and the harm she suffers.

6.3 Objectification, Stereotyping and Scientific Self-Knowledge

Having identified third-stage objectification as an interpretive moral wrong, and having used the examples of sexual objectification and commodification to draw out some of the challenges associated with adequately theorising such wrongs, I now wish to consider a case that brings the issue of the radicality of theorising such wrongs into a sharper contemporary focus. This case not only problematises the relationship between common sense and interpretive moral wrongs, it also has a bearing on the relationship between distinctively scientific self-knowledge and interpretive moral wrongs.

Of the interpretive moral wrongs that we have considered so far (discrimination, stereotyping, and objectification), it is fair to say that the primary relevance of discrimination lies in straightforwardly political contexts, where concerns about alleged injustice are uppermost. The relevance of objectification by contrast, and in particular third-stage objectification, may well seem more obscure. The examples we have considered suggest that concerns about objectification may be relevant to research ethics, to the ethics of care work, and to sexual politics. But the relevance of concerns about objectification is by no means confined to such cases. Rather they are also relevant to any area connected with the production and application of the types of knowledge associated with what Foucault termed the 'human' sciences. Any area, that is, in which human subjects are made the objects of authoritative knowledge – especially if there is a simultaneous risk of the instrumentalisation of human beings in the application of such knowledge. The example of commodification has already hinted at some of the issues here, since the classical economic theories with which Marx was concerned occupy an area of overlap between science and common sense. The example I now wish to turn to involves another area in which the relationship between scientific self-knowledge and common sense has become problematic: the science of human genetics.

6.3.1 Objectification in Genetic Research

Recent attempts to compile human genetic databases, combining genetic, medical and genealogical information, have focused on relatively discrete groups, such

as the population of Iceland. The reasons given for picking out groups like the Icelanders have to do with the availability of historical and contemporary records, and the history of isolation and population stability, which have made the Icelanders an unusually genetically homogenous group. (Though some commentators have emphasised that the Icelanders' degree of genetic homogeneity is often exaggerated – many straightforwardly ideological elements, with a long and distinctive history of their own, have tended to strongly influence the decision to focus on the Icelanders as a peculiarly valuable genetic resource.[12])

The idea that the Icelandic population might be exploited as a resource for genetic research has proved controversial, and it certainly has some disquieting aspects, sufficient to give rise to concerns about the objectification of its subjects. But it is less easy to pin down the real source of such concerns, so as to assess their reasonableness. It could be argued that as biomedical research goes, the establishment of genetic databases is relatively innocuous – being substantially free of any risk of direct physical harm to the research subjects. Moreover, in line with the prevailing tendency in the ethics of biomedical research to equate objectification with simple instrumentalisation, it might well be though that it would be a straightforward matter to safeguard against the objectification of the Icelanders, via suitable consent arrangements.

Would this be sufficient though to allay the concerns of anyone seriously worried about the objectification of the Icelandic population via genetic research? If the only basis for such concern was the danger that the research population would be treated purely instrumentally, it might well be thought that there can be no remaining grounds for serious concern. But matters are not as simple as that. At a 2004 conference on the ethics of genetic databases an Icelandic speaker expressed the concern that, over and above any problems relating to establishing suitable consent arrangements, there is something potentially disquieting in the very idea that the Icelandic population might be used in the proposed way, as a genetic model.[13] This concern was questioned by an audience member. What, the questioner wanted to know, is so ethically disquieting about serving as a model? Granted that to serve as a model is to serve in some sense as a means (it was argued), so long as those involved consent to this treatment, and are not reduced to the status of a mere means, no significant ethical concerns arise. Here the discussion reached an impasse. The speaker evidently felt that the questioner had not fully appreciated the distinctive ethical and political challenges that arise in connection with using populations as genetic models. But the questioner seemed entirely satisfied that, worries about straightforward instrumentalisation having been assuaged, there is no reason why serving as a genetic model should generate any disquiet. We can however pursue this question here: just what are the specific ethical concerns (if any) raised by serving as a scientific model? The answer to this question, I want to suggest, is that individuals who serve as scientific 'models' are, in principle at least, exposed to a form of third-stage

---

[12]See Árnason 2004.

[13]The speaker was Gardar Árnason – see Árnason 2004.

objectification – whether or not they are also in danger of being reduced to the status of a mere means in any more straightforward way.

To see why this is so we need to return to the issue of ideological stereotyping, discussed in Chapter 4. Our concern with ideological stereotyping focused on two main types of case: Firstly, the case in which a possible categorisation-generalisation complex turns out to altogether lack predictive and explanatory power, despite the fact that the relevant generalisations are true, both in general, and in the particular case(s) under consideration. The example given was that of the set of people with the same number of letters in their first and last names. Even if it should turn out to be the case that certain generalisations are true of the members of this group both at the general level, and in particular cases, it seems overwhelmingly likely that such an arbitrary categorisation-generalisation complex would, in any real case, altogether lack predictive and explanatory power. Secondly, there was the case in which generalisations based on other possible categorisation-generalisation complexes turn out to be true, and yield predictive and explanatory power, but only under 'distorting' socio-historical conditions. The example given here was that of a specific ethnic or religious group whose members turned out to have a characteristic propensity for risk-taking – perhaps as a result of certain historical factors, such as a history of exploitation and oppression. Such categorisation-generalisation complexes could turn out to have genuine predictive and explanatory power – albeit only as a consequence of the relevant regrettable historical factors. I argued earlier that anti-discrimination safeguards quite properly require us to put facts associated with categorisation-generalisation complexes of either of the above kinds 'out of play', when making decisions that directly affect the social and economic fortunes of individuals (such as job selection, and insurance offers etc.). By contrast with the above cases of potential ideological stereotyping, we can also identify, notionally at least, a third type of categorisation-generalisation complex, which would capture certain essential truths about the members of a group, that reflect their true nature. Marx's contention that humans are distinguished by their capacity for genuinely creative and productive activity would, supposing it were true, be an element in a complex of this kind. In such cases we could say that the relevant generalisations yield genuine predictive and explanatory power, and do not reflect any form of historical distortion.

On the basis of this threefold distinction between the different ways in which the generalisations associated with a given categorisation-generalisation complex might have come to be true of the members of a group, we can distinguish three different senses in which a generalisation or generalisations might be said to be true of them. Firstly, the generalisation might be true *merely accidentally* – as it would be *merely accidentally* true of the set of people with the same number of letters in their first and last names, that they had a propensity to risk-taking. Secondly, the generalisation might be true *as a matter of historical contingency* – as it would be true as a matter of historical contingency that a given ethnic or religious group had a propensity to risk-taking, in any case in which this feature was a product of a regrettable history of exploitation and oppression. Thirdly, a generalisation might be *essentially true* of an individual or group, in a case in which the relevant generalisation yielded genuine

predictive and explanatory power, and was not associated with any 'distorting' historical conditions.[14]

With this terminology in place, we can capture in a more satisfactory way what it is that is disquieting about serving as a model in a scientific study. One aspect of the problem of ideological stereotyping that we have thus far avoided confronting explicitly is that of the relative scientificity of categorisation-generalisation complexes: how far, and on what grounds, can such complexes be regarded as genuinely scientific? Turning now to address this question, we find that it is hard to give a clear answer. I have already made considerable use of the idea that a given categorisation-generalisation complex may or may not yield predictive and explanatory power. The majority of commentators have tended to view these two qualities as the most important hallmarks of a genuinely scientific theory.[15] And with these criteria in mind, such obviously arbitrary categorisations as 'people whose first and last names have the same number of letters' evidently lack scientific legitimacy.[16] A genuinely scientific study does not aim to arrive at merely *accidentally* true generalisations. But if yielding predictive and explanatory power is indeed the proper test by which to distinguish genuinely scientific from unscientific categorisation-generalisation complexes, there seems nothing to prevent generalisations that are true merely as a matter of historical contingency being regarded as genuinely scientific. Earlier we saw that such generalisations are a fertile basis for ideological stereotypes. As such, they are quite properly put out of play in contexts in which potential discrimination is a concern. Thus if it is correct to say that some ideological stereotypes may yield predictive and explanatory power, it seems necessary to concede that some scientific categorisation-generalisation complexes may also be ideological stereotypes.

It follows that, while it is certainly possible to employ the criteria of predictive and explanatory power to rule out some wilder-seeming possibilities (such as the category 'persons with equal numbers of letters in their first and last names'); and while these same criteria may conceivably also serve to rule out certain other

---

[14]Note that while it might be tempting to say that in the first two of these cases the generalisation is contingently true of the members of the group, while in the third case it is necessarily true, there are good reasons to prefer the terminology I have chosen. To speak of generalisations as contingently true in the first two types of case would risk obscuring the important differences between them, since some generalisations that would be contingently true in this sense would possess genuine predictive and explanatory power, while others would not. Furthermore, to speak of generalisations that are essentially true, in the sense explained above, as being necessarily true, would risk obscuring the important differences between statements that are true in virtue of their meaning, and generalisations that are true in virtue of the underlying nature of the entities over which they generalise.

[15]See for example Newton-Smith 1981, p. 4.

[16]That said, it is an alarming feature of contemporary popular presentations of science that they frequently present observed statistical correlations as scientific findings, without providing any convincing causal-explanatory underpinning. As pointed out earlier, until we have some idea why members of a particular group might be more likely to display a particular feature, we know nothing reliably about their true propensity to display that feature, let alone whether anything related to their membership of the relevant group is likely to explain that propensity.

possible categorisation-generalisation complexes (such as those associated with a discredited concept such as 'race') they will not serve to distinguish categorisation-generalisation complexes whose predictive and explanatory power is a consequence of historical distortion from categorisation-generalisation complexes that manage to capture the true nature of the group in question. A further dimension of the difficulty here comes to light when we consider the evident scientific value of concepts that *do* seem to reflect a regrettable history of exploitation and oppression. Consider for example a historically central social-scientific concept such as (socio-economic) class. The concept of social class is widely held to have a legitimate and useful role to play in social science research and theory. The concept may be usefully invoked in various political, economic and sociological contexts for a range of predictive and explanatory purposes. But since, as we have seen, the history of any group (thus of any social class) might (arguably, *must*) involve elements of damage and distortion, there seems nothing in principle to prevent a concept such as social class combining predictive and explanatory power with playing a role as an ideological stereotype.[17]

The fact that under present conditions, and all foreseeable future conditions, the concepts available to the human sciences stand under suspicion, at least, of possessing predictive and explanatory power only as ideological stereotypes, perhaps explains the disquiet that serving as a scientific model or specimen sometimes seems to elicit. If we simply ask: 'what is wrong with serving as a model or specimen?' we may well conclude, if we restrict ourselves to thinking of the subjects of such research as primarily at risk of instrumentalisation, that there is nothing wrong with doing so – provided suitable consent is obtained. But the above discussion of science and ideological stereotyping has suggested that the third-stage objectification of research subjects may still be a danger, even where the subject(s) in question are not otherwise harmed or wronged by the relevant procedure.

In addition to serving as a means, the subject of scientific research serves as a model/specimen. That is, she serves as an exemplar of some categorised-and-generalised-about aspect or subset of humanity. (If this were not the case, no valid scientific generalisations could be forthcoming from the study.) However, as a model or specimen, she is exposed to objectification in a sense that goes beyond any danger of simple instrumentalisation. She is both treated as a means (not, ideally, as a mere means), and simultaneously treated as a model or exemplar of what *may* turn out to be an ideological stereotype. The source of the concern is not simply that what is found to be true of the group as a whole is likely also to be regarded as true of the individual. On the one hand, what is true of the group may turn out not to be true of the individual at all. On the other hand, supposing what is true of the group is true of the individual, it may be true of the individual merely accidentally, and this may be immediately apparent to all concerned. The more worrying case is one in which what is true of the group is also true of the individual, but is true of her

---

[17]It is worth noting that while it might be considered legitimate in practice to use the concept of social class to predict and explain e.g. my political or economic behaviour, the same concept could not be used by e.g. an insurer to estimate my propensity for risk-taking without raising concerns about discrimination.

only as a matter of historical contingency. In this case it is plausible to think that the study might yield scientifically well-founded conclusions, and enable researchers to generate findings with real scientific value. However, the corresponding danger would lie in thinking that, since the study has generated scientifically valid results, it cannot at the same time be perpetuating ideological stereotypes. For, as we have seen, in any case in which what turns out to be true of an individual is true only as a matter of historical contingency, we may have results that are predictively and explanatorily valuable, but still rest upon an ideological stereotype. What is true of an individual or group purely as a matter of historical contingency might, owing to its apparently respectable scientific credentials, come to be regarded as reflecting the underlying nature of the relevant individual or group.

It turns out then that some potentially quite reasonable concerns attach to cases in which individuals or groups come to play the role of scientific models or specimens. Consider again, for purposes of comparison, the victim of sexual objectification. It might conceivably be true that, due to a legacy of historical oppression, a number of claims concerning 'what women want' (e.g. to be subservient) will turn out to be true of many or most women. In a case like this, the women of whom the relevant generalisations are, as a matter of historical contingency, true, may well come to believe that they are *essentially* true. In such a case they may try to ensure that their behaviour conforms to these generalisations – they will actively seek to 'pass' as women, thus conceived. However, from the fact that the stereotype is true of them, and is also predictively and explanatorily valuable (thus qualifying as scientific, on one common construal), it will not follow that it expresses any essential truth about them. For, if things had been different, what is now true of them might well not have been true. However, in prevailing circumstances, and given the cultural power of the stereotype, it is unlikely that the individuals concerned will be particularly alert to the distinction between what is true of them solely as a matter of historical contingency, and what might turn out to be true of them essentially. What they *are* will likely seem to be straightforwardly a matter of what is true of them. Or, where it is acknowledged that what is true of them might include generalisations that are true *merely accidentally*, a matter of what is true of them in such a way as to yield meaningful predictive and explanatory power (thus, on some interpretations at least, a matter of what is *scientifically* true).

Similarly, in some cases serving as a scientific model or specimen could encourage those involved to regard the resulting scientific account of what they are as expressing essential truths about them, when it is in reality an ideological stereotype. For the research subject to take up a particular scientific concept – such as the concept of social class – into her own self-understanding may lead to her trimming herself to fit a damaging and distorting mould. More specifically, the 'model' who features in *genetic* research, as an instance of a general category, which may be associated with a categorisation-generalisation complex that is in reality an ideological stereotype, is in great danger of interpreting herself, and being interpreted by others, in a stereotyped manner. What is scientifically true of her will, in the circumstances, be likely to be taken to express what she is – since contemporary ideology has it that genetics tells us what we really are, as no other science can.

The concern here is not simply that no study can complete the complicated story of what any human individual is. It is not that the story must be incomplete, but rather that *it must be distorting, where the object of study is distorted*. The mythology has it that genetic databases tell us something about the nature of the relevant group, independent of social and historical factors. And when biomedical science wishes to sample a 'pure' and undistorted human population the preferred candidates are groups like the Icelanders. But the Icelandic population can no more be held up as an example of undistorted humanity than can any other human population.[18] The reasons why people settled in Iceland initially are reasonably well-understood, and are saturated with social factors. Similar stories may be told about the 'bottlenecks' through which the population has since passed. Whatever predictive and explanatory power may result from a present-day enquiry into the genetic makeup of the Icelanders, it cannot reasonably be claimed to be proof against ideological stereotyping, either of the Icelanders or of others.

Of course the entire point of such exercises may be to discover something about the legacy of certain social-historical distortions – for example if the research concerns the effects of certain culturally-linked diseases (perhaps reinforced through sex-selection). In such cases the research subject would have to accept that while the resulting categorisation-generalisation complex might well be true of her, and possess genuine predictive and explanatory power, it will not necessarily reflect what she is 'by nature'. In general, the category of 'human' is no more guaranteed to be proof against historical distortion than the category of 'working class'. Both will yield predictive and explanatory power in particular socio-historical settings, and both may represent ideological stereotypes when set against what human beings might have been had things been different.

However, when the goal of the research is not that of discovering specific truths concerning a distinctive group, and the factors that have caused them to be as they are, but is instead that of uncovering some exemplar of pure and undistorted humanity (or womanhood etc.), the danger of third-stage objectification is much more pressing. To be placed in a context in which one serves as a means by serving as a model or specimen, is to be peculiarly exposed to the combination of wrongs that we have found to be associated with objectification, on more radical interpretations of that concept. The subject of scientific research is exposed to ideological self-stereotyping, in addition to any instrumentalisation she may suffer, because the whole basis of her inclusion in the study is that she is to be identified with the group under investigation *as it has historically turned out* – any conception of which must, as we have seen, at least stand under suspicion of being an ideological stereotype. We may conclude that, given that all categorisation-generalisation complexes pertaining to human beings must at least stand under suspicion of being vehicles for (possibly beneficent) ideological stereotyping, there is a permanent threat of wrong to the participants in such research, which is substantially independent of any threat to their well-being, or their capacity for self-determination.

---

[18] Árnason 2004, p. 34.

6.4 Interpretive Moral Wrongs and Human Dignity

My intention in the foregoing has been to draw out some of the more controversial-seeming aspects of the analysis of interpretive moral wrongs to highlight the extent of the challenge that the theorising of such wrongs poses for applied and professional ethics. I now want to consider what normative resources remain available to us, in the face of this challenge.

If we ask wherein the wrongness of instrumentalising or exploitative treatment lies, on a traditional Kantian account, the answer seems to be that, ultimately, it is wrong because it is incompatible with the dignity of a human person to be treated in such a way.[19] A significant feature of this type of response is that it gives us a way to say that it is morally problematic to be the victim of such treatment even in cases where the victim does not seem to suffer any identifiable harm. On this basis we can insist that both relatively benign forms of slavery, and (more relevant for contemporary applied and professional ethics) beneficent paternalism, are morally problematic. At the same time, the Kantian response to instrumentalisation suggests that it is incompatible with the dignity of the agent to treat another in such a way. Any and every human person has a stake in the idea of human dignity, and if human dignity is diminished in any of its persons, we all suffer some form of diminishment as a consequence (which, again, may not register as a recognisable harm).[20]

An illuminating contrast with utilitarianism is available here. The claim that it is incompatible with my own dignity to humiliate another, given that human dignity is something in which we all share, seems far more plausible than the claim that it is incompatible with my own happiness to make others miserable – many historical examples could be cited in support of the contention that a good deal of human happiness has been purchased at the price of making others miserable. We cannot therefore be claimed to have a collective stake in human happiness as plausibly as we can be claimed to have a collective stake in human dignity.

However, what the Kantian account does not do is exhaust the potential of this appeal to dignity. The Kantian concern with dignity amounts to a concern with the dignity of human beings as (rational) *agents*. It is held to be incompatible with the dignity of a human agent to suffer systematic frustration of her choices, or to systematically frustrate the choices of another. But if no more than this is said, the concern with dignity is in danger of collapsing into a concern with anti-instrumentalisation – nothing more. (For example, the kind of concern with instrumentalism that we find in the 'four principles' approach, where respecting the autonomy of others is interpreted as a matter of respecting their freedom of choice and action.)

The foregoing analysis of interpretive moral wrongs throws the limitations of such an approach into relief. In an earlier chapter I made reference to Peter Singer's

---

[19]Kant 1996, pp. 83–5.

[20]Of course, Kant would not say that the moral wrongness of a failure to respect the dignity of another resides in the consequent diminishment of the idea of human dignity in general. His concern is with the rational inconsistency involved in any such act, not with its actual or potential consequences. But that is not the same as denying that it *has* significant general consequences.

concept of speciesism – being the arbitrary preference for members of one's own species in moral matters.[21] Prima facie, speciesism, as defined by Singer, is ethically unjustifiable. But for all its superficial parallels with racism and sexism, it is hard to see how, given the analysis of discrimination developed above, speciesism could, like racism and sexism, lie at the root of a genuinely discriminatory social practice. To operate a double-standard in ethical, political or economic decision-making would in some cases be speciesist. And in at least some cases such activity would technically fit the description of discrimination as outlined above (as a form of 'rigging against. . .' with all its necessary interpretive and ideological accompaniments). Why then might we hesitate to regard a speciesist action as also potentially discriminatory? I would contend that it is because, in the ordinary run of things, the victims of such speciesism will not be ontological persons, equipped with a sophisticated capacity for self-interpretation. No speciesist act, whose victim was a member of some relatively unsophisticated non-human mammalian species, could have the consequences of discrimination, or indeed of any other interpretive moral wrong. The significance of wrongs such as sexual objectification, whose victims are the sorts of beings who can maintain a sophisticated self-conception, is that they make available a conception of the victim that she is able to incorporate into her self-understanding. It follows that interpretive moral wrongs proper must have ontological persons as their victims (though for most purposes it is unnecessary to make this explicit).

The idea that human dignity is connected with our capacity to act on the basis of freely chosen rational principles, rather than simply being tugged this way and that by desire and inclination, is well-recognised in ethics. But the idea that human dignity is intimately connected with self-knowledge has an even longer history. While the command of the oracle of Apollo to 'know thyself' may have slipped from its former prominence in the western moral consciousness, it is widely acknowledged that there is something troubling to both the popular and the theoretical moral consciousness about the idea that e.g. a flourishing human life could be lived in conditions of systematic ignorance or deception. For example, to bring up a child in complete ignorance of human suffering would be to compromise her dignity, to some extent, even if she was happier as a result. To bring up the same child in ignorance of her own mortality intensifies the affront, even if, again, the consequences for her measurable state of happiness are all positive. When pushed to explain why we feel it is so important to know the truth, even when the truth is unpalatable, we naturally turn to the idea that there is something about living in a state of ignorance or deception that is incompatible with human dignity. And if human dignity is compromised where we are deceived about the nature of the outside world, the affront to our dignity must be that much greater if it is not simply our knowledge but our self-knowledge that is restricted or compromised.

If, as I am suggesting, to be in a state of ignorance or deception concerning one's own nature involves an affront to one's human dignity, independent of any further

---

[21] See above, Section 2.3.2.

consequences it may have for one's happiness, or capacity to act, then the moral significance of a principle of respect for persons is not exhausted by a concern with the status of persons as sentient and self-determining agents. I fail to show you the respect due to another person if I behave so as to cause you to have a distorted or inadequate self-conception, or if I act in a manner that ensures that you maintain such a self-conception when I am in a position to remedy it, irrespective of any further implications my conduct may have for your future options for action.

It follows that the failure to respect another as an *end* may involve more than a failure to respect her as an *agent*. The undermining or restriction of another's self-conception is intrinsically ethically significant, and it is significant because persons are more than simply agents. The dignity of persons rests in part on their capacity for sophisticated self-awareness and self-knowledge. By restricting or compromising another's capacity for self-knowledge (otherwise put, by inducing 'false consciousness'), we wrong them as persons in a manner that goes beyond any wrong associated with the undermining or restriction of their capacity for autonomous agency.

I do not mean to suggest that such ignorance or 'false consciousness' is morally significant only if deliberately induced. In Chapter 3, I claimed that discrimination can be either deliberate, or 'quasi-deliberate' – in cases in which the discriminating agent has a clear duty not to discriminate, whether acted on or not. With respect to many examples of objectification, and many, though not all, of the restrictions and distortions of self-understanding that are consequent on discrimination and stereotyping, it will no doubt be unrealistic to suggest that they are deliberately induced, in any meaningful sense. My concern is that a distorted self-understanding is always ethically undesirable, and, where induced deliberately, or quasi-deliberately, culpably so.

The challenge that confronts contemporary applied and professional ethics, in the face of the various interpretive moral wrongs we have been considering, is thus as follows: We must come to appreciate the potential impacts of ideological stereotypes in our dealings with ontological persons. Such stereotypes play a key part in the interpretive moral wrongs of discrimination, objectification and commodification. In these cases they may play a legitimating ideological role, affecting the interpretation of the victim by the perpetrator, and by interested third parties, though without necessarily affecting the victim's self-interpretation. They may also play a second type of ideological role, in which they affect the self-conscious goals of the victim, with indirect consequences for her dignity, through causing her to act 'voluntarily' in pursuit of what are essentially alien goals. Finally, they may play a third type of ideological role, in which they directly subvert the victim's self-understanding, with the above-mentioned indirect implications for her dignity, and also with direct implications for her dignity, inasmuch as being blessed with, and being permitted to maintain, an undistorted self-conception, is a key component in the dignity of persons.

References

Árnason, Gardar. 2004. Interbreeding within the Icelandic population is high compared with mice or fruit-flies. In *Blood and data*, eds. G. Árnason et al., 29–35. Reykjavik: University of Iceland Press.

Kant, Immanuel. 1996. Groundwork of the metaphysics of morals. In *Immanuel Kant: practical philosophy*, ed. Mary Gregor, 41–108. Cambridge: Cambridge University Press.

Marx, Karl. 1954. *Capital*, Vol. 1 (trans: Moore, S. and Aveling, E.). London: Lawrence and Wishart.

Marx, Karl. 1967. Economic and philosophic manuscripts (1844). In *Writings of the young Marx on philosophy and society*, eds. Loyd D. Easton and Kurt H. Guddat, 283–337. New York, NY: Anchor Doubleday.

Newton-Smith, W.H. 1981. *The rationality of science.* London: Routledge.

Nussbaum, Martha C. 1995. Objectification. *Philosophy and Public Affairs* 24(4):249–291.

Part II
Sources and Foundations

Chapter 7
Hegel and Recognition

Abstract It would be surprising if the central finding of Part I of this book – that respecting the dignity of ontological persons requires that we respect their capacity for undistorted self-interpretation, in addition to respecting their capacity for autonomous action – lay behind the widespread but somewhat inchoate concern with such interpretive moral wrongs as objectification, yet had gone entirely unnoticed by previous enquirers. The primary concern of Part II is to round out and contextualise the findings of Part I by tracing the tradition of post-Kantian ethical reflection that can be said to be centrally concerned with the ethics of self-knowledge. In doing so it will suggest some significant modifications to the account of the normative foundations of the ethics of self-knowledge set out in Chapter 6. The process begins in this chapter with an exploration of Hegel's view of the nature and moral importance of the recognition of others, as an element in our own self-realisation.

7.1 Recognition

We saw in Chapter 2 that the principle of respect for persons requires us never to reduce a person to the status of a mere means. The continuing ethical relevance of this principle can be questioned on a range of grounds, as discussed above. Nevertheless, I wish to claim that, suitably revised, it still has a significant and distinctive role to play. To this end I have analysed a specific class of moral wrongs, which I refer to as 'interpretive' moral wrongs. These have the distinctive feature that they embody an interpretation of their victim(s). When one is discriminated against, stereotyped, or objectified, one is discriminated against (etc.) 'as' a member of a particular group. The interpretation of the victim as a member of the relevant group, with all that entails, is built into the offence. (By contrast, when one is simply (e.g.) bullied or exploited, one is not bullied or exploited 'as' anyone in particular.)

Focusing on objectification, as an example of an interpretive moral wrong, we saw that in third-stage objectification (and such related phenomena as commodification, and the objectification of scientific models/specimens), the fact that it embodies an interpretation of the victim has particular significance. The taking up of an ideological stereotype into the self-interpretation of the victim brings about her reduction to the status of a means in a manner that goes beyond the more familiar reduction associated with mere instrumentalisation. This form of reduction to the status of a

means is substantially unacknowledged in more conventional interpretations of the idea of respect for persons.

Once this mechanism is seen at work in objectification, it is easier to appreciate the distinctive ethical issues raised by other interpretive moral wrongs. Even where the particular self-interpretation taken up by the victim consists mainly or exclusively of truths, it may still amount to an ideological stereotype – a distorted conception of what she really is. By being taken up in this way the distortions of the stereotype are reinforced and potentially compounded. The interpretation ceases to be something that resides primarily in the eye of the beholder. Instead it becomes something with which the victim can actively play along – thereby reinforcing its apparent authority.

Whatever other applications the principle of respect for persons may have in contemporary applied and professional ethics, it retains a key role in relation to interpretive moral wrongs. It is particularly in relation to conduct that affects others at the level of their self-interpretation that we need, and will continue to need, a principle of respect for persons in private and professional life. This concern with ideological stereotyping, and consequent distorting self-interpretation, is not however easily accommodated within mainstream contemporary ethics. Modern moral philosophy has concentrated largely on the rightness and wrongness of certain *acts*.[1] And even when, as in contemporary neo-Aristotelianism, the focus shifts from acts to the qualities of agents, there remains a tendency to portray virtues as derivative of right action, or at least as explicable in terms of rules or principles of right action.[2] The example of these interpretive moral wrongs suggests that this emphasis is mistaken. Not simply how we act toward others but how we interpret them should be a major concern of ethics. We have duties not to ideologically stereotype others, which are to some extent independent of duties concerning how we otherwise treat them. Of course, the two sets of duties are not entirely separate. One reason why the instrumental stereotyping of others is wrong is that it is likely to encourage further instrumentalising treatment. But it does not follow that such stereotyping is not also wrong in itself, independent of its implications for overt behaviour. As our examples of commodification and of the scientific 'model' show, there are potential wrongs involved in ideological stereotyping that are independent of any wrongs (or harms) the victims might suffer from any associated instrumentalising treatment. But are these duties not ultimately derivative of a duty not to instrumentalise others, as Kantian ethics would suggest? At the close of the last chapter I argued that the duty not to induce or maintain distorted or otherwise inadequate self-conceptions is not derivative of a duty not to (simply) instrumentalise, but that both duties share a common root in the duty to respect human dignity. In the remainder of this study I shall argue that it is also possible to identify an independent foundation for the first of these duties in the virtue of *honesty*. Showing how this is the case will involve

---

[1] See Anscombe 1997.

[2] See for example Hursthouse 1999, chapter 1.

exploring a tradition of philosophical ethical reflection that runs slightly to one side of the modern mainstream.

The ethical implications of self-interpretation are, as I have suggested, of no more than passing concern for the mainstream of modern moral philosophy. However, the ethical significance of self-interpretation was explicitly emphasised by the philosopher who has probably been Kant's most influential critic: G.W.F. Hegel. For Hegel, and for the tradition of ethical reflection he initiated, the themes of self-interpretation and *recognition* are key concerns of ethics.[3]

The idea that we are subject to a key ethical imperative to *recognise* ontological persons as free rational beings, capable of sophisticated self-interpretation, and not properly reducible to the status of a mere means even in thought, is a central feature of Hegel's moral philosophy.[4] Hegel considers the basis of this imperative to lie in our inherent drive to achieve a state that he refers to as 'self-certainty'. By self-certainty Hegel means in effect a state of *conscious self-possession*.[5] For a being, such as an ordinary human being, who is capable of self-interpretation, the healthy development of the personality will be in the direction of, and must involve, such a state of conscious self-possession.

The ideal of self-certainty is closely linked to an ideal that we have already seen to be key to the ethics of respect for persons: that of self-determination. But at the same time it involves a significant expansion of that ideal. As we have seen, it is possible to construe self-determination as if to be self-determining were simply a matter of active choice-making, and the deliberate pursuit of our own goals. The individual who is self-determining in this sense is simply one who makes decisions for herself, rather than allowing them to be made by others on her behalf. A richer conception of self-determination is based on an ideal of autonomy. To be truly self-determining, in *this* sense, is to engage in *principled* choice-making. One makes the choices one does from obedience to certain rules or imperatives (ideally of course, on a Kantian view, from obedience to the categorical imperative), rather than on the basis of (e.g.) evanescent desires. However, it is possible to envisage a still more expansive conception of self-determination. As previously noted (Section 2.3.2), in order to obey a rule I must understand myself as an addressee of the rule. All rule-following of the relevant sort therefore requires having a self-conception, of some kind. A truly self-determining choice would thus be not simply a choice originating from the individual concerned, nor simply a choice made on the basis of a self-given rule, but, in addition, a choice made in light of an adequate conception of who and what one is – and with a view to what is most important, given who and what one is. (Which is perhaps just to say that having the capacity to act on the basis of an adequate self-conception is a frequently overlooked aspect of any really robust conception of autonomy.)

---

[3]See Wood 1990, chapter 4 and passim.

[4]Ibid.

[5]Ibid., p. 84.

So, the idea that human beings are subject to an inherent drive to achieve conscious self-possession seems to receive support from the ethics of respect for persons, inasmuch as such an ethic is built around an ideal of self-determination; and self-determination, properly understood, presupposes sophisticated self-interpretation. Showing due respect for self-determining beings will accordingly require that we show suitable respect for their attempts to achieve adequate self-understanding: conscious self-possession, or 'self-certainty'. Expressed negatively, such a state of self-possession is equivalent to a state of non-alienation – the over-coming of a fundamental self-estrangement. But how, and under what conditions, might such self-certainty be achieved? The key feature of Hegel's analysis of the preconditions for self-certainty is that it involves the possibility of experiencing oneself as an *object*.

Self-certainty involves self-consciousness, and self-consciousness is a paradoxi-cal phenomenon. Consciousness is relational. To say that x is conscious is to imply that there is some y (which may be nothing more than an idea) such that x is con-scious of y. Ordinarily then, we think of consciousness as a relation involving at least two terms, a subject term and an object term. But self-consciousness must then be a peculiar sort of relation, since it would be an instance of self-relation – in this case x and y, subject and object, would coincide. Paradoxical though this idea of self-consciousness as self-relation may be, it is at the same time a feature of the most ordinary of experiences. Whenever we have a belief or desire we are typically aware not only of the belief or desire itself, but also of the fact that the relevant belief or idea is our own – that we have that belief or desire. Thus to have a belief or desire is simultaneously to be aware of ourselves as the 'bearer' of that belief or desire. To that extent self-consciousness belongs inescapably to the fabric of our day-to-day lives. Of course it may be that in some cases we fall into self-forgetfulness, in which self-awareness recedes. Perhaps into fits of 'abstraction' – as in T. S. Eliot's example of 'music that is heard so deeply that you are the music, while the music lasts'.[6] These would presumably be cases in which we are aware of an 'object' (the music) but no longer aware of ourselves as aware of that object – a state of consciousness without self-consciousness. But such experiences are note-worthy in part because they are atypical. They stand out against the background of everyday existence, in which awareness and self-awareness, consciousness and self-consciousness, constantly accompany each other.

Nevertheless, if we are to be self-aware at all we must be capable of grasping our-selves as an object of awareness. How is this achievable? According to Hegel, the quest for self-certainty begins with desire, and the appropriation of things. We aim to understand ourselves through the objects we desire, and strive to possess.[7] This strategy proves insufficient however, because it involves an unsatisfactory depen-dence on mere objects. Straightforward desire-satisfaction does not go to confirm our sense of ourselves, or lead to self-possession or self-certainty. Instead it tends

---

[6]Eliot 1974.

[7]Hegel 1977, pp. 104–6. See also Wood 1990, pp. 84–5.

to lead to a dissipation of self, in a world of things (in effect, we 'lose ourselves' in the world).[8] Establishing and maintaining an adequate sense of self thus requires an encounter with an object in which we are able to find ourselves, rather than losing ourselves – an object in which we discover ourselves. This in turn requires a special sort of object. According to Hegel, it demands a self-negating object. That is, an object that announces inescapably that it is other than, more than, the mere 'thing' that it is.[9]

Hegel's conception of this process of self-discovery through an encounter with a self-negating object can helpfully be thought of in the following way: when the desire for self-certainty tries to find satisfaction in an inanimate thing it accomplishes only its own dissipation. We can perfectly well act upon and acquire objects on the basis of our desires. But they 'reflect' nothing back to us. Crucially, they are not objects that offer us a perspective on ourselves. They are not objects for which we, in our turn, are objects. When we encounter another subject however, the situation is altered fundamentally. For now we are in the presence of an object for which we are ourselves an object. This 'other' therefore holds what we need to satisfy our desire for self-certainty, because she offers us what no other object can: a perspective on ourselves. Only through contact with another consciousness (which for most of us begins at an earlier point in our lives than we can possibly recall) can we achieve self-consciousness – a sense of ourselves as a possible other, to another. Hegel concludes:

> Self-consciousness achieves its satisfaction only in another self-consciousness.[10]

The important implication of this is that I am not self-sufficient in my quest for self-certainty. In order to achieve freedom as self-possession I am fundamentally dependent on my interactions with others.

7.1.1 Hegel on Master and Slave

The fact that self-consciousness needs another self-consciousness generates an important vulnerability. If two beings meet, both in search of self-certainty, each holds the resources the other needs to complete his quest. But while I need the 'recognition' that the other can give me, I initially have no particular investment in giving him the recognition he needs. Indeed, as long as the other exists I am vulnerable, to the extent that I am dependent on another for the satisfaction of my most fundamental need.[11] There is potential for serious conflict here. Each party will have a motive to extort the recognition he requires from the other, while simultaneously

---

[8]Hegel 1977, p. 106; Wood 1990, pp. 84–5. It seems that the appropriation of everyday things would terminate in a condition where we 'gain the world, only to lose ourselves'.

[9]Wood 1990, p. 85.

[10]Hegel 1977, p. 110.

[11]Wood 1990, p. 86.

attempting to distance himself as far as possible from the troubling dependency that this involves. Hegel goes so far as to claim that in this encounter each party will attempt to annihilate the other. The realisation of their mutual dependency will precipitate a fight to the death.[12]

A temporary armistice will be achieved, Hegel argues, when one party elects, from fear of death, to play a subservient role. In exchange for simply being allowed to live, he will provide the other with the recognition he demands, while accepting that he will not receive similar recognition in his turn. Henceforth, one party will be master, the other slave.[13] It may seem that there is something perverse in the attempt to annihilate the very being who is able to provide the external perspective we crave. Would this not ultimately be self-defeating? It seems that it would, if slavery amounted to complete annihilation of personhood. But in fact the master does not demand that the slave be entirely reduced to the status of a thing. Rather, the slave must become a (conscious) instrument in the master's hands. The slave must labour on the master's behalf, in service of the master's goals, while continuing to provide the master, partly by this means, with the recognition he requires.[14]

The peace thereby achieved is not permanently stable however. The roles of master and slave contain the seeds of their own destruction, because, it turns out, the condition of slavery is actually more propitious for the development of genuine self-certainty than is that of mastery. Through being required to work for the master's goals, the slave is able to achieve a unique perspective on his own goals.[15] He is forced to defer gratification of many of his day-to-day desires. And as he learns what it is to work for goals that are not linked to his own immediate desires, he will learn to distance himself from those desires, and to discipline himself in respect of them. At the same time, the slave's day-to-day labour will involve experience of the recalcitrance of physical things. As he struggles to master them on his master's behalf, the effort will leave its mark on him. But at the same time, it will ultimately be the slave's mark, and not that of the master, that is left on the things with which he works. Through the process of labour, then, it will be the slave, and not the master, who most profoundly transforms and appropriates his physical surroundings.[16]

In this process the slave will acquire a measure of self-sufficiency and self-restraint. Paradoxically, slavery will teach him self-mastery. More and more he will have to modify his fundamental desire for egoistic self-certainty. He will ultimately learn that adequate recognition must stem from a relationship between equals. True self-certainty does not involve simply extorting recognition from another. This is not because, or not simply because, to be worth anything the required recognition

---

[12]Hegel 1977, pp. 113–4.

[13]Ibid., p. 115.

[14]Ibid., pp. 114–8. Hegel also observes that the deeper meaning of the struggle to the death is not to annihilate the other, but to risk one's own life (ibid., pp. 113–4). In gambling with our own destruction we demonstrate (to ourselves and to the other) our non-identity with the determinate physical particular we also are.

[15]Ibid., pp. 117–9. See also Wood 1990, pp. 86–8.

[16]Hegel 1977, pp. 118–9.

must be freely given (we can only get true recognition and self-certainty from an encounter with a being we regard as free).[17] It is also, more profoundly, because true recognition involves an element of genuine *reflection*. It involves *finding and recognising oneself in another*. In order genuinely to find oneself reflected in another one must attain a perspective that is potentially universal – a perspective that is potentially shareable with 'others like me'. Only if what one looks for in others by way of recognition is something that the other can authentically share can one actually recognise oneself in the other.[18]

The situation of the slave is propitious then not simply for the achievement of recognition but for the achievement of recognition in its most authentic form. Servitude will ultimately open the door to *universal self-consciousness*. The slave's progress toward self-certainty will not have been governed simply by the forms of recognition that are actually open to him, but also by what he has learned along the way concerning what it is possible to look for, and find, in others. The slave achieves the capacity to truly see himself in another, because he learns to look for what can be genuinely universal. There is then a creative dimension to this whole process.[19] Through the struggle for recognition, the slave ultimately arrives at a position that is propitious for self-certainty. But what the slave has learned has not come wholly and directly from that struggle. The slave has become acquainted with a higher truth concerning himself. He has learned that *he is not what he is*, or not simply what he is (that is, a slave, whose *raison d'être* is to labour to satisfy the master's desires). He has learned that he has the potential for free rational action. And this will suggest to him that, essentially, he is not a slave at all. Servitude has given him insight into his own higher possibilities. The servant learns the hard way that:

> [W]hat is important for self-worth is not the gratification of desire but the dignity of formally free agency.[20]

Through servitude then, the slave acquires a type of self-certainty that is inaccessible to the master. Servitude serves as a preparation for life in a community of free rational persons.[21] By contrast, the master's experience of exercising arbitrary power, his physical and psychic dependency on the slave, and habituation to instant and effortless gratification, lead to a lack of self-mastery. The role of master produces the opposite of what was initially sought – the master sought self-possession, but the role of master undermines genuine self-possession.[22] Mastery might look like freedom, but it is not true freedom, since when the master looks on others all he

---

[17] Wood 1990, p. 89.

[18] Ibid., pp. 88–92.

[19] As Wood (1990, p. 92) expresses it: 'a self actualises itself when it makes itself into what it needs to be in order to satisfy its desire for self-certainty'.

[20] Ibid., p. 89.

[21] Ibid., p. 88. See also Hegel 1977, pp. 119–20.

[22] Strikingly illustrated in the film *Gosford Park* (Robert Altman, 2001) a central theme of which is the servant's ability to anticipate her employer's desires, to the point where she knows what he wants before he knows himself.

will see reflected is his own narrow ego, with its selfish and parochial desires. The master cannot find himself in another, because all that can be reflected back to him is his own particularity, his own limitedness.[23]

Nevertheless, to properly realise the state of self-certainty, the slave still needs to experience actual recognition. The understanding he has achieved of his higher possibilities does not make him entirely self-sufficient. He may have learned something about genuine recognition, but he still needs to experience such recognition. The master-slave phase is an important staging post on the journey, but is not the end of the journey. Genuine self-certainty will only be achieved by escaping the condition of servitude. The slave needs to find or establish a community within which true recognition – founded on universal self-consciousness – is more than just a theoretical possibility.[24]

7.2 Dignity and Universal Self-Consciousness

In Hegel's view it is only in the midst of an actual community, in which genuine mutual recognition in line with the higher possibility of universal self-consciousness can be manifested, that the full dignity of the human being will be realised and respected.

On a Kantian interpretation of the dignity of persons, our dignity is intimately connected with our capacity for self-determination. Only those with such a capacity possess dignity in the relevant sense. As we have seen however, the concept of self-determination is open to more than one interpretation. If self-determination is interpreted as individual decision-making, and dignity in turn interpreted as derived from, or at least intimately linked with, the capacity for self-determination, we risk reducing the idea of dignity to absurdity. Referring back to Regan's claim (Section 2.3.2) that non-human subjects-of-a-life command our respect just as human subjects do, and for similar reasons, such a view would invite the conclusion that the dignity of human persons is nothing more than the dignity that attaches to any and every choice-making subject of a life. But while it may be that we should bring non-human subjects of a life into the moral fold – not simply in the manner advocated by Singer, where their interests are taken into account, but in the manner advocated by Regan, in which beings with projects of their own are protected from instrumentalisation and excessive exploitation – to suggest that their status as moral patients confers on them the dignity we associate with human persons seems to risk depriving the notion of dignity of all meaningful ethical content. While it may be that both human and non-human subjects of a life deserve protection from exploitation and purely instrumental treatment, the idea that the dignity of human persons resides in a capacity they share with farm animals seems a reductio of the very idea of human dignity.

---

[23]Wood 1990, p. 89.

[24]Ibid., p. 93. Hegel 1977, p. 290.

It would be mere speciesism, though, to insist without further explanation that a capacity that confers dignity on human beings fails to confer dignity on those non-humans that possess it. What is it then about the sort of self-determination that is distinctive of human beings that confers dignity on those who possess it, in a manner that the self-determination of the non-human subjects of lives that Regan aims to protect does not?

Kant, as we have seen, conceives of the moral community as a kingdom of ends.[25] In the Kantian kingdom of ends there is no distinction between moral patients and moral agents. All potential members of a kingdom of ends are subject to the moral law, and thus are both moral patients and moral agents. All therefore possess what we might term 'full' moral personhood. In earlier chapters we saw that this all-or-nothing conception of moral personhood has become increasingly untenable over recent decades. Singer, Regan and others have advanced powerful arguments to the effect that some non-moral-agents are nevertheless moral patients; and as a consequence the traditional conception of moral personhood has become fragmented. The moral community is evidently wider than the community of moral agents.

However, it need not be concluded that Kant's conception of the kingdom of ends is henceforth redundant. The concept of a community of moral persons bound by reciprocal duties of respect and non-interference might still legitimately be used to characterise the moral situation of human beings, and any non-humans that might be discovered to have a capacity for the sort of sophisticated self-determination – 'autonomy' – that is of primary interest for Kant. It would simply have to be acknowledged that such a 'kingdom of ends' would not be the entire moral community – that beyond its borders there would lie a community of other moral persons, who possess moral patiency, but not moral agency. On this basis it would be possible to argue that a non-speciesist construal of the dignity of the potential members of a kingdom of ends is possible, since the members of *this* moral community would be, uniquely, full moral persons. Dignity in the relevant sense would attach to those who are both moral agents and moral patients, but not to those who are simply moral patients.

The problem with this line of argument though is that it is hard to see how the possession of moral agency could make *that* sort of difference. Why should an important moral quality such as dignity attach to full moral persons, rather than attaching to all moral patients equally? Why should moral agency confer a unique type of dignity? As long as we think of potential members of a kingdom of ends along the lines we have been exploring in this part of the discussion, these questions seem hard to answer. A moral person who is both the locus of obligations and the subject of obligations doesn't seem to be so very special, in comparison with those who are simply the loci of obligations. And as long as we think of the members of a potential kingdom of ends as simply bound by mutual obligations it seems hopeless to try to spell out wherein their unique dignity consists.

---

[25] See above Section 2.2, and Kant 1996, pp. 83–5.

But Hegel's view of a moral community as bound not only by mutual obligations, but by the sort of mutual recognition that universal self-consciousness makes possible, casts an altogether more promising light on the idea of the unique moral significance of potential members of a kingdom of ends. The conclusion seems to me to be inescapable that it is not full moral personhood per se that confers dignity, but the capacity to *recognise* oneself and others as members of a potential kingdom of ends. That is, dignity does not attach to 'full' moral persons, simply as moral agents – tracking their capacity to act on the basis of principles such as Kant's formula of humanity. Instead, our dignity attaches to our capacity to recognise each other, and thereby understand ourselves, as potential members of such a community.

Such a community would appear very like Kant's kingdom of ends, but with a particular distinctive feature. The dignity of the potential members of this community would be bound up with the fact that they reciprocally recognise each other as 'ends' in a manner that goes beyond implicitly acknowledging, in their actions, that their fellows are independent beings who are pursuing projects of their own, and have a right not to see those projects arbitrarily frustrated. The only kind of recognition that seems to be demanded of the Kantian moral agent is recognition of the factors that contribute to the moral patiency of others. Paradoxically, it seems, precisely what Kantian ethics fails to explain is why moral agents, as opposed to mere moral patients, are entitled to any special kind of consideration. It is precisely the applicability of the Kantian model to non-human subjects of lives that threatens to subvert the traditional idea of the kingdom of ends as a morally significant community. And it is no solution to argue that the potential members of this community are possessed of dignity, where mere moral patients are not, since we will then want to know what is morally significant about dignity, thus understood.

In the image of the Hegelian subject, in search of self-certainty, and thereby required to recognise others as members of a community of moral agents, we see an example of the explicit recognition of others as self-interpreting beings. And once we appreciate that there may be something like a kingdom of ends, which is founded not simply on mutual obligations, but more profoundly on mutual recognition, the basis of the special dignity of the members of such a community becomes apparent. Their capacity for self-understanding and mutual recognition confers on them a unique dignity, and simultaneously exposes them to unique harms and wrongs – they can, for example, be the victims of interpretive moral wrongs, as other moral patients cannot. To think that members of a potential kingdom of ends based not simply on 'autonomy' and reciprocal obligations, but on recognition, possess a dignity, and an associated entitlement to special forms of respect and consideration, that is not possessed by other moral patients, is not necessarily an arbitrary prejudice then. However, this dignity, and those entitlements, are not grounded merely in their capacity for autonomy. They are grounded in their capacity for sophisticated self-interpretation (which no doubt in turn partly grounds their capacity for moral agency).

Put that way, one might be forgiven for concluding that these two qualities will always be found together in any case, and on some such basis it could be argued that all of this is already implicit in Kant's conception of what it is to be a potential

member of a kingdom of ends. But if this is the case, it is surely important to make this feature explicit. Moreover, it is also important to point out what follows from my more general argument: that if this is so we have duties to potential members of a kingdom of ends that go beyond the duty not to instrumentalise or exploit them (specifically, we have a duty not to subject them to interpretive moral wrongs). I shall not attempt to establish whether the Kantian moral agent must in fact be understood to possess, of necessity, the relevant capacity for recognition and sophisticated self-interpretation; but if she must, then it is evident that her human dignity will rest partly on her possession of that capacity. This being so, such dignity is in danger of being compromised not simply by instrumental treatment, but by being forced to live under conditions that inhibit genuine self-understanding.

To genuinely respect another person involves, in addition to refraining from instrumentalisation, showing consideration for the fact that she is the sort of being the adequacy and accuracy of whose self-conception is hugely important *for her*. Concomitantly, any ethical approach that restricts itself to the avoidance of instrumentalism (whether in the name of respect for persons or not), or, worse, restricts itself to the maximisation of general happiness, overlooks an important aspect of human personhood. The importance of this aspect of the dignity of persons – the aspect that relates to self-knowledge rather than directly to autonomous agency – is attested to in the seriousness with which we regard such wrongs as discrimination, stereotyping, and objectification. The human victims of classic discriminatory syndromes such as racism and sexism will be predisposed to suffering a distinctive wrong, which is additional to any associated harms they might suffer. Discrimination is not, as we have seen, reducible to any form of simple 'injustice'. It is injustice 'with an account' – interpretive injustice, with a 'story' or ideological rationale attached. And human beings, being the reflective beings they are – the subject-objects of self-understanding and self-knowledge – may suffer wrongs through being discriminated against in addition to the harms to which they are exposed as a consequence of their sentience and agency.

To be the victim of simple injustice is not necessarily disempowering. Indeed, the reverse may be the case. By appreciating the injustice of one's situation, and fighting against it, one may gain a stronger and more adequate sense of one's own identity (and this may be accompanied by the development of solidarity in a group). But to model all forms of injustice on this sort of case would be to underestimate how subtle and pernicious some forms of injustice can be. Where injustice takes the form of discrimination, there is the constant danger that rather than being outraged and energised by one's experience of injustice, one will tend increasingly to acquiesce in it – the accompanying rationale having become lodged parasite-like in one's own self-conception. Members of the discriminated-against group will be in danger of coming to understand themselves as they are understood by their oppressors, feeling that the treatment they receive is permissible, even merited. This makes discrimination a particularly hard-to-challenge form of injustice. Controlling how the victims of injustice are able to think of themselves is an efficient means by which to intensify and perpetuate injustice. But, as I have suggested, the ethical significance of discrimination is not exhausted by its efficacy as a means. To cause someone to

hold a damaged and damaging self-conception is to injure their dignity, to wrong them, irrespective of any associated harms that may or may not follow.

The great merit of Hegel's discussion of the importance of self-certainty, and the conditions under which the required universal self-consciousness (which amounts to the full *recognition* of others as ends in themselves) is possible, is that it encourages us to take seriously this idea that we have specific duties to others as self-interpreting beings. The Hegelian 'history' of the origins of self-consciousness, and the basis of moral recognition, thus goes some way to explaining the ethical significance of interpretive moral wrongs.

7.3 Essentialism and Political Liberalism

In both its ethical and its theoretical aspects, the philosophy of Hegel is built around a concept of self-actualisation.[26] Self-actualisation is a teleological notion. To actualise one's self in the relevant sense is to actualise one's telos, to 'become who you are', or, in a slightly less paradoxical formulation, 'become who you really are'.[27] One of the attractions of an ethic of self-actualisation is that it can comfortably accommodate both the idea of morally significant harms, and the idea of moral wrongs. Failure to achieve self-actualisation can take a purely physical form. The self-actualisation of humans or other animals implies a particular type of physical existence – having certain basic needs met, being able to engage in certain forms of physical behaviour.[28] Thus a physically damaged or distorted existence falls short of self-actualisation; and causing another to live in circumstances of significant physical deprivation is interpretable as harming them. But failure to achieve self-actualisation can equally involve certain sorts of psychic deprivation, which cannot be reduced to straightforward 'harm'. We have already seen, for example, how ideological stereotypes encourage individuals to interpret themselves in terms of concepts that reflect historical distortions. Thus understood, stereotyping is inimical to self-actualisation. When I am led to base my self-understanding on an ideological stereotype I fall short of self-actualisation as surely as when I am deprived of the physical basis for a flourishing existence. In this sort of case however, it may be more appropriate to say that I am wronged, than that I am harmed – as in cases of

---

[26]Wood 1990, pp. 30–5.

[27]Ibid. Wood emphasises that Hegel's theory is not teleological in the sense that e.g. utilitarianism is a teleological theory (i.e. in the sense that it commences with and revolves around a specific conception of the good). But it is certainly teleological in the sense in which Aristotle's moral theory is teleological, and this is the sense I have in mind here. Aristotle's theory is teleological inasmuch as his specification of the good as happiness is in terms of an identity to be actualised. Hegel's theory similarly commences with 'the conception of a certain self or identity to be exercised or actualised, to be embodied or expressed in action' (Wood 1990, p. 31).

[28]What neo-Aristotelians term 'external' goods. See Macintyre 1981, chapter 14. Marx's views on the indispensability of an adequate supply of external goods for a flourishing human existence are thoroughly Aristotelian in character. See Marx 1967, p. 293.

paternalism, it may be that decisions that are properly those of the client/patient her-self are pre-empted, in a manner that is otherwise beneficial for her. It is not that the paternalistic medic, or the well-meaning social scientist, necessarily misconstrues the other's well-being, causing her more harm than good – rather, the point is that each may wrong her even while successfully promoting her good.

A key objection to the ethics of self-actualisation, though, is that it is *essentialistic*. It understands the process of becoming what one is as one of 'coming into' or 'actualising' one's essence. Essentialism has a long and illustrious philo-sophical history. But at the same time it remains highly controversial, particularly in the English-speaking philosophical world. The idea that humanity has an essence, which is in principle distinct from its observable nature under prevailing social and historical conditions, has been thought by many to be unacceptably metaphysical.[29] Here, though, it is necessary to acknowledge that I have already made use of a num-ber of essentialist ideas in the course of the last few chapters. My discussion to this point has made crucial use of two essentialistic distinctions: the distinction between what is true of a group as a matter of historical contingency, and what is essen-tially true of that group; and the distinction between a prevailing 'distorted' form of existence and an underlying undistorted 'nature' or 'essence' of the group. These distinctions are connected, and the connection between them is evident when we bear in mind that it is with reference to the nature or essence of the group that we are able to select, among the various generalisations that are true of them, those that are essentially true of them. Thus it is the nature or essence of the group that is the more fundamental notion.

The idea that humanity has an essence, which may or may not be successfully actualised in practice, does indeed open the door to a range of metaphysical ques-tions. Initially, it may seem no more mysterious or counter-intuitive than the idea that non-human species such as (e.g.) tigers have an essence or nature, which may or may not be actualised, depending on the circumstances. The idea that tigers have a nature, which cannot be properly actualised in, say, the artificial conditions of a zoo, is neither novel nor particularly challenging to grasp. (The idea that 'you can't really see a tiger in a zoo' is familiar enough, and makes a perfectly sensible point.) There is then nothing essentially mysterious about some versions of the appeal to the essence or nature of a species. What is more likely to give reasonable people trouble is the thought that essentialist notions can be transferred from the context of (e.g.) natural history to that of ethical and political thought. The plausibility of the idea that the tiger has an essence, which cannot be properly actualised in captivity, rests largely on the evident significance, in this type of case, of the contrast between the artificial conditions of a zoo, and conditions in the wild. But the latter distinc-tion has no meaningful parallel in the case of human beings. While it is easy enough

---

[29]Modern Anglophone opposition to metaphysical essentialism begins with Locke, who rejects the notion of 'substantial forms'. See Locke 1975, book 3, chapter 6, sections 2–10. Lockean 'real' essences are not essences in the relevant sense, as they belong to the physical constitution of entities. The anti-essentialist tradition continues through Hume to such twentieth-century figures as Wittgenstein, Quine and Rorty.

to distinguish natural from artificial environments in relation to tigers, it is impossible to do the same in relation to human beings. This is partly because examples of 'natural' humanity are not to be found. But, more fundamentally, it is because all talk of human nature seems to invite the paradoxical conclusion that humans are essentially unnatural: that it belongs to the nature of humans that they inhabit substantially artificial physical and cultural environments.

The project of developing a philosophy of self-actualisation in respect of human beings must therefore contend with the difficulty that if there is such a thing as the human essence, it is not something that we can directly observe 'in the field', as it were. Any ethic of self-actualisation applied to human beings must somehow accommodate the fact that human nature, if observable at all, can only be studied as it exhibits itself under artificial conditions. Hegel's essentialism takes account of this problem. In his view, a scientific approach to human self-actualisation is only possible by taking what would now be termed a 'hermeneutic' approach to understanding what it is to be human. The guiding idea behind such an approach is that what it is to be human will be evident not in the way that humans appear to a detached observer, but in the various ways in which they have interpreted themselves. The empirical basis that Hegel's theory of human self-actualisation requires is therefore drawn from the study of history.

History records the successive arising and passing away of a variety of forms of human life and culture. But history is not simply a record of events, and the growth and decay of human cultural forms. The point of pursuing an enquiry into the human essence and human self-actualisation historically is that history provides a record of the ways in which those events and cultural forms were understood by the human agents who lived through them.[30] History does not give us the past. It cannot be a detached approach to the study of an obscure 'object'. (It is not a branch of natural history.) In giving us an insight into the sense that human beings have made of the events they have lived through it constitutes a record of the unfolding of human self-consciousness itself.[31] This unfolding can be traced in the history of philosophy. But it is also evident in the history of other cultural activities, such as art, religion, politics and government etc. For Hegel, the philosophical study of the unfolding of human self-consciousness in history reveals that the historical growth and decay of various cultural forms is not, as might otherwise be thought, an aimless, directionless process, but (egregious episodes of decadence and reversal notwithstanding) a process of progress and development. This process is a teleological one, tending toward a quite specific goal (that history took the particular course it did is held to be a contingent matter, but that it tends toward the goal in question is not). It is practically impossible to imagine a philosophical, artistic, religious or political outlook that did not at the same time embody a more or less adequate view of what it is to be a human being. Human history records a great drama of alienation

---

[30]See Houlgate 1991, chapter 1.

[31]Ibid. See also Hegel 1975, p. 101.

(self-estrangement) and (ultimately) self-actualisation, which amounts to an evolution of human self-consciousness.

In this idea that history records a process of evolution toward full human self-consciousness and self-possession, a number of significant assumptions are made concerning what it is to be human. Some of these assumptions – for example those connected with the idea that the unfolding of human self-consciousness is an essentially progressive process – seem contestable. (Though Hegel would argue that the study of history lends firm support to them.) Others – for example that humans are in fact self-interpreting beings, and that their self-interpretations are important to them – seem more or less indisputable. (At least, it is hard to imagine having anything further to say to someone who was prepared to deny them). But of course the two sets of assumptions are not unconnected. If it is true that human beings are in fact self-interpreting beings, and that their self-interpretations are important to them; and if it is true that history is composed of a series of successive human self-interpretations; then later interpretations will typically have been informed by earlier ones, and the general Hegelian teleological picture becomes hard to avoid. At the level of events, the case for moral or indeed any other form of teleological development might be hard to make out, particularly in the face of the major conflicts of the last century. But at the level of human self-consciousness, and in the face of our tendency to interpret the entirety of the past as leading up to ourselves, it is hard to avoid the conclusion that as time passes we become increasingly well-equipped to know ourselves.

However attractive such hermeneutic essentialism may be though, it remains at odds with dominant contemporary values.[32] We cannot claim that we don't know roughly what would be involved in bringing a moral community of the kind envisaged by Kant and Hegel into existence. But equally, it cannot be claimed that such an existence is a live possibility for most of us. Of course, the dignity of the members of such a community is held not to lie in the actual conditions under which they live and labour, but in the fact that recognition and universal self-consciousness are possible for them – their dignity resides not in what they actually are, but in what it is possible for them to become. Such ideas are a familiar enough feature of political debate, and are not easily dismissed as mere rhetoric.[33] But the idea that our self-actualisation depends on actually bringing some such community into existence has proved to be deeply offensive to liberal political tastes.

The defining characteristic of the political liberal is the conviction that the only grounds on which we are justified in restricting individual liberty are those of maximising liberty, or preventing harm.[34] Consequently, the idea that we might be required to restrict liberty in the service of a moral ideal such as universal

---

[32] See Wood 1990, Conclusion.

[33] Such political milestones as Martin Luther King's 'I have a dream' speech testify to their perennial appeal.

[34] See Mill 1962, p. 193.

self-consciousness, is considered by liberals to be politically noxious.[35] Of course, the fact that Hegel's ideal is itself an ideal of liberty could reasonably be expected to make some difference here. Hegel, it could be said, simply envisages the establishment of a moral community in which genuine freedom would become a reality. But whatever else may be said for or against this ideal, it is an ideal that articulates a specific view of what it is to be human – a specific view of the essence and the self-actualisation of human beings. And this is what makes it so offensive to liberals. For many liberals Hegel seems a dangerous ideologue, the friend of totalitarianism. And, for some commentators, the close association of Hegel's ideas with those of Marx lends plausibility to this view.[36]

In the wake of the bloody twentieth century then, Hegel's High Enlightenment idealism tends to look excessively prescriptive; and, insofar as it relies on a form of essentialism, dangerously irrationalistic. Such essentialistic views, it has been claimed, install a particular ideal as a universal goal, without properly subjecting this ideal itself to rational criticism. In the view of prominent anti-essentialists, the development of political institutions is properly a matter of trial and error, rather than teleological self-actualisation.[37] It is legitimate to intervene in this process rationally, through 'piecemeal' social engineering. But it is only legitimate to do so by applying the type of rationality we find in science, which eschews dogma and speculation, in favour of controlled experimentation, and the strenuous testing of falsifiable conjectures.[38] One powerful motive behind this sort of liberal approach it that it avoids broaching potentially interminable discussions concerning competing conceptions of human nature. As such it is part and parcel of a consciously anti-essentialist stance that pervades liberal approaches to ethics and political philosophy.[39] This anti-essentialistic stance of modern liberal thought is significant for my analysis for two reasons.

Firstly, having shown that the general significance of interpretive moral wrongs can be explained on the basis of a Hegelian ethic of self-actualisation and recognition, it is apparent why our broadly liberal political and philosophical culture might

---

[35] See Berlin 1958, pp. 17–8.

[36] Probably the most influential and outspoken recent critic of Hegel, who sees an intimate relationship between Hegel's ideas and the excesses of institutionalised Marxism, is Karl Popper (see Popper 1945).

[37] Ibid.

[38] I should note that while I think there are good reasons not to be an essentialist, I don't endorse the anti-essentialist arguments of Popper. I mention Popper's views here because at this point I am concerned with the de facto political and cultural acceptability of essentialism, rather than its philosophical defensibility.

[39] See for example Rorty 1989. Whether more thoroughgoing liberal political philosophies, such as that of Kant, can ultimately avoid entanglement in essentialist ideas of human nature is a much-disputed matter. The categorical imperative, for example, avoids mentioning any particular human ends, but the Kantian explanation of its importance cannot avoid doing so – minimally, through the idea of the good will. And of course this problem is not unconnected with that of whether membership of the kingdom of ends requires us merely to practice non-interference with others, or requires positive Hegelian 'recognition' among its members.

have shied away from discussing offences such as objectification, stereotyping and discrimination in great detail. If ideological stereotyping can be understood along essentialistic lines, by contrasting what individuals and groups have been permitted to become under prevailing socio-historical conditions, with what they would have been had they been allowed to develop freely in line with their underlying nature, then it is also a concept that liberal ethical and political philosophies will tend to cut themselves off from, by their own anti-essentialist stance.[40] Secondly, it suggests that any contemporary ethic of respect for persons that takes interpretive moral wrongs more seriously is likely to encounter considerable resistance from liberals.

In light of the above points, the present analysis faces a twofold difficulty: it must avoid the problems consequent on adopting a liberal anti-essentialist approach, which tends to cut itself off from an adequate understanding of the nature and importance of ideological stereotyping; but at the same time it will have to confront the perceived problems facing forms of essentialism.

In the previous chapter I distinguished a radical from a reformist response to understanding interpretive moral wrongs. I argued that the example of sexual objectification in particular shows us that the reformist approach seems bound to misinterpret, and consequently to misrepresent, the challenges associated with combating such wrongs. I also said that the task of theorising interpretive moral wrongs demands a radical approach (and accordingly cannot necessarily be expected to deliver easy solutions with immediate common sense appeal). In endorsing the radical approach I may seem to have embraced some form of essentialism. (For example, the sort of combined essentialism and radicalism found in Marx, and exemplified in his analysis of commodification.) However, in my view there is enough in the contemporary critique of essentialism to give us pause. In a sense, the problem with essentialist forms of radicalism is that the answers they provide are bought too cheaply. Many Marxists have stressed the scientific basis of Marxism, and it is not uncommon to combine Marxism and scientific realism to produce a view that is both radical and essentialist. But, as discussed in Chapter 7, we also have reason to be suspicious of at least some scientific accounts of human nature and society. To the extent that these suspicions are well-founded, such scientifically-based forms of essentialism might turn out to be much less radical than they take themselves to be – insofar as they fail to allow for the fact that some scientific accounts of

---

[40]This sheds interesting light on the finding that the principle of respect for persons has fallen from prominence during the latter part of the twentieth century. For the same period is one in which radical and potentially all-embracing forms of ideology critique, such as Marxism and feminism (so-called 'metanarratives') have declined in cultural influence, while unorthodox applications of well-established liberal ethical principles – such as Singer's application of a utilitarian model to our relations with non-humans – have achieved quietly revolutionary results. The tempting conclusion is that a robust principle of respect for persons always required a metaphysical underpinning, even if it has, until relatively recently, been able to distance itself from essentialism without any obviously damaging consequences.

human nature may perpetuate ideological stereotypes irrespective of any predictive and explanatory power they might possess.

It would be a great pity though if the attempt to establish adequate theoretical foundations for a principled opposition to ideological stereotyping, and the offences against persons that involve such stereotyping, were to be abandoned in the face of contemporary qualms about essentialistic views of human nature. To avoid this outcome, what seems to be required is some further development of the key Hegelian insights concerning the ethical importance of recognition and self-interpretation, which does not depend on controversial appeals to an essentialistic conception of human nature. In short, we seem to need to develop a radical but non-essentialistic philosophy of recognition. Fortunately, the work of a series of twentieth-century thinkers in the Hegelian tradition provides the resources such a philosophy needs. This tradition developed increasingly fine-grained analyses of the ways in which the various forms of self-understanding that are available to us, or are foisted upon us, are associated with distorted and/or debasing self-conceptions. It will be the task of the next and succeeding chapters, then, to show how, on the basis of this tradition, a genuinely radical but nevertheless non-essentialistic ethic of self-interpretation might be developed.

References

Anscombe, Elizabeth. 1997. Modern moral philosophy. In *Virtue ethics*, eds. R. Crisp and M. Slote, 26–44. Oxford: Oxford University Press.

Berlin, Isaiah. 1958. *The concept of liberty.* Oxford: Clarendon Press.

Eliot, T.S. 1974. The dry salvages. In *T.S. Eliot collected poems 1909–1962*. London: Faber and Faber.

Hegel, G.W.F. 1975. *Lectures on the philosophy of world history* (trans: Nisbet, H.B.). Cambridge: Cambridge University Press.

Hegel, G.W.F. 1977. *Phenomenology of spirit* (trans: Miller, A.V.). Oxford: Oxford University Press.

Houlgate, Stephen. 1991. *Freedom, truth and history: An introduction to Hegel's philosophy.* London: Routledge.

Hursthouse, Rosalind. 1999. *On virtue ethics.* Oxford: Oxford University Press.

Kant, Immanuel. 1996. Groundwork of the metaphysics of morals. In *Immanuel Kant: practical philosophy*, ed. Mary Gregor, 41–108. Cambridge: Cambridge University Press.

Locke, John. 1975. *An essay concerning human understanding*, ed. Peter H. Nidditch. Oxford: Oxford University Press.

Macintyre, Alasdair. 1981. *After virtue.* London: Duckworth.

Marx, Karl. 1967. Economic and Philosophic Manuscripts (1844). In *Writings of the young Marx on philosophy and society*, eds. Loyd D. Easton and Kurt H. Guddat, 283–337. New York, NY: Anchor Doubleday.

Mill, John Stuart. 1962. On liberty. In *Utilitarianism*, ed. Mary Warnock, 126–250. London: Fontana.

Popper, Karl. 1945. *The open society and its enemies*, Vol. 2. London: Routledge and Kegan Paul.

Rorty, Richard. 1989. *Contingency, irony and solidarity.* Cambridge: Cambridge University Press.

Wood, Allen W. 1990. *Hegel's ethical thought.* Cambridge: Cambridge University Press.

Chapter 8
Heidegger and Authenticity

Abstract The Hegelian concept of recognition ties the ethics of self-interpretation to an ethic of self-realisation. However, it rests on essentialist claims about the nature and possibility of self-knowledge that are unlikely to be sympathetically viewed in the generally liberal and anti-essentialist atmosphere of contemporary applied and professional ethics. This chapter aims to show how a non-essentialistic alternative might be developed, which preserves some key features of Hegel's model, on the basis of Heidegger's account of authenticity. Heidegger can be said to take a *sceptical essentialist* view of human nature. Just as knowledge remains an issue for the epistemological sceptic, even while he rejects it, so Heidegger's approach to the question of what it is to be human acknowledges that our own being never ceases to be an issue for us, while recognising that to embrace any of the (ultimately instrumentally derived) self-conceptions that are available to us would be inauthentic.

8.1 Liberalism, Essentialism and Positivism

Key resources for developing the theoretical foundations of a radical and principled opposition to ideological stereotyping can be extracted from the philosophy of Martin Heidegger – specifically, from Heidegger's concept of authenticity. To show how this can be done will be the task of the present chapter. However, it is important to acknowledge at the outset that the ethical implications of Heidegger's thought are notoriously difficult to extract. They tend to emerge only as secondary implications of his ontological enquiries.[1] In order to properly grasp the potential of a Heideggerian opposition to ideological stereotyping, it is necessary to briefly consider what is at stake in the choice between essentialism and *positivism*.

   Liberal anti-essentialism is often associated with some form of positivism. 'Positivism' is a term with a wide array of (interrelated) meanings, and what exactly it should be taken to mean in any given case tends to depend heavily on the precise context. Common to all forms of positivism though are an opposition to 'metaphysics', and a suspicion of all forms of speculation concerning unobservable

---

[1]Cooper 1990, pp. 34–6.

phenomena, in science and elsewhere. Since the essences with which essentialists are concerned are usually considered to be prime examples of unobservable metaphysical phenomena, positivism and essentialism are incompatible.[2]

One important role for claims about essences is to underpin and explain certain counterfactual claims we might wish to make about individuals or kinds. For example, in trying to shed light on the nature of the historical forces at work in a given period, a historian might want to consider how events would have turned out if a particular influential individual had been brought up in a different religion. This sort of investigation can only proceed on the assumption that it is possible to distinguish between the essential and the accidental properties of the individual concerned. If it makes sense to say that the very same individual could have been brought up in a different religion, then this can only be because his having been brought up in the religion he was in fact brought up in was an accidental feature of him – a feature he could have lacked while remaining essentially the same individual. It is only because we are comfortable with the idea that such features are accidental features of individuals that we regard such speculation as coherent, and potentially informative. If, by contrast, someone had wished to speculate on how things would have turned out if the self-same individual had been born to different parents, we might well reject such speculation as meaningless. It seems unreasonable to think that the very same individual could have been born to different parents, for a child with different parents would surely not have been the same child. (In light of surrogate parenthood we should probably be more precise and say something about the origins of the gametes involved – but the general point still holds.) This sort of view assumes that being born to the particular people who were in fact one's parents must be an essential feature of an individual.[3]

Counterfactual claims concerning species or kinds can give rise to similar questions, and can be treated in a broadly similar way. So, for example, it seems perfectly coherent to wonder, in the context of an anthropological enquiry, how things might have turned out if early humans had never spread beyond Africa. (Since it is presumably an accidental feature of early humans that they did in fact spread beyond Africa.) On the other hand, the question 'how would things have turned out if humans had evolved for a purely aquatic existence?' might well be considered meaningless, it being regarded as essential to humans that they are terrestrial mammals.

In picking out the essential properties of an individual or species, we will in some cases be picking out the defining properties of the being in question. Historically, the role of essences in underpinning and explaining claims concerning definitions has been at least as important as their role in underpinning and explaining counterfactual claims about individuals and kinds. Essences have also been invoked to explain claims made about such things as abstract virtues.[4] These various types of claims

---

[2]For a useful discussion see Hacking 1983, chapter 3. See also Popper 1963.

[3]Kripke 1980, pp. 111–5.

[4]The Socratic appeal to forms to explain claims about e.g. piety is a familiar example of this kind of appeal. See for example Plato 1954.

are more or less inescapable in philosophy, and elsewhere. (It is hard to imagine how we could proceed with a discourse that had entirely dispensed with counter-factual claims, and claims concerning the defining characteristics of such things as abstract virtues.) And in such cases it is necessary to invoke the idea of an essence because essential properties cannot plausibly be presented as *self-evidently* essential. Controversies concerning e.g. the defining features of justice, or human nature, are practically interminable because what it is for an act or institution to be just, or a being to be human, is not at all self-evident.

Positivistic approaches to knowledge tend to brush such questions aside however. Questions about essences are dismissed as eccentric and superfluous, and it is taken to be obvious to any serious enquirer what the essential features of an entity are (or, if not obvious, then at least amenable to empirical scientific enquiry).[5] This strategy can be persuasive in some areas of the natural sciences, where, due to the degree of specialised expertise involved, we are usually prepared to operate what Hilary Putnam refers to as a '*division of linguistic labour*'.[6] We don't tend to argue with respected physicists about the essential properties of water. Instead we usually accept the considered views of specialists. And if there is controversy amongst specialists, it is usually confined to the professional arena – it rarely has an impact on the wider public consciousness.[7]

However, in areas where there is less automatic deference to experts, such positivistic briskness doesn't work nearly so well. In respect of questions concerning human nature for example, we acknowledge no natural authorities. Indeed, one significant problem with modern scientific attempts to make (e.g. genetically-based) declarations concerning various aspects of human nature is that – presumably in order to avoid provoking instant incredulity – they tend to eschew the sort of scientific rhetoric that implies privileged insight, unavailable to the layperson, in favour of conceptions of human nature that effectively endorse common sense views. As a consequence, such claims may do little more than endow common sense opinions with a spurious scientific authority.[8]

My enquiry to this point has suggested that without a broadly essentialistic conception of human nature, which embraces an idea of how individuals could or would have been, but for socio-historical distortions, we cannot make sense of the idea of ideological stereotyping that is central to the revised approach to the ethics of

---

[5]See for example Popper 1945, pp. 9–21.

[6]Putnam 1993, p. 155.

[7]This is not to say that all such disputes are always scientifically resolvable, or that they never have an impact outside of academic circles. Kuhn (1962) discusses a number of historical cases in which a dispute about definitions turned out to be irresolvable, and was ended only by a 'paradigm shift' in the relevant science, which was at least as much a social and psychological shift as a scientific one. A more recent illustration is provided by the disagreements among astronomers during the summer of 2006 concerning the definition of a planet. These had the effect of pushing the number of recognised planets in the solar system from the traditional 9, to 12, and then down to 8, all within the space of a couple of weeks.

[8]This is a core theme of Husserl 1970a. It is picked up and developed by the Critical Theorists of the Frankfurt School, in particular by Marcuse and Habermas.

respect for persons that I am proposing. That said however, it has to be conceded that in the light of the liberal criticisms outlined in the previous chapter, it is hard to make essentialism look ethically and politically attractive. Thus the choice between essentialism and positivism may look to be a choice between two unattractive alternatives. There is however a third option. This third option arises from, but is not reducible to, a phenomenological form of essentialism.

8.2 Phenomenological Essentialism

A general philosophical concern with essentialism has been a feature of intellectual life in Europe for two and a half millennia, on and off. Edmund Husserl, who founded the modern phenomenological movement at the turn of the twentieth-century, regarded some form of essentialism as necessary to any properly grounded philosophical approach. Husserl viewed the rise of positivism, particularly in psychology, as a precursor of irrationalism – both in science, and, ultimately, throughout society as a whole.[9] At the same time however, Husserl was conscious of the shortcomings of traditional forms of essentialism – the unattractiveness of which partly explained the popularity of scientific positivism as an alternative.[10] He therefore sought to develop an alternative philosophical approach to the study of essences, which would proceed via the systematic exploration of essential structures of human consciousness.

Husserl did not consider himself to have been the first to approach this task in such a way. He acknowledges that both Descartes and Kant made very important contributions toward the same goal.[11] However, he also believed that his predecessors had made significant errors along the way. While acknowledging his debt to Kant, and in particular Kant's achievement in isolating the categorial elements presupposed in all judgements of experience, Husserl felt that Kant had fallen into errors of emphasis, and also of method.

With respect to his method, Husserl criticises Kant for the 'constructive' quality of his approach. No one was a more vigorous opponent of metaphysical speculation than Kant. And yet he himself relied on a 'constructively inferring' method that presupposed many of the results it set out to obtain, particularly in relation to the fundamentals of logic. Kant's error here, Husserl argues, is to overlook the indispensability of an 'intuitively disclosing' method in philosophy. It is not enough to find oneself inexorably driven to certain conclusions by established logical principles. In such fundamental matters it is necessary rather to (mentally) 'see' the phenomena with which the study aims to deal. [12] This criticism of Kant is implicit in

---

[9] Husserl 1970a, part I, section 2, and passim.

[10] Ibid., part I, section 5.

[11] Ibid., part II, sections 16–21; part III A, sections 28–32. See also Husserl 1995, Introduction and First Meditation.

[12] For Husserl's criticisms of Kant see in particular Husserl 1970a, part III A, section 30; 1995, p. 86.

the phenomenological slogan, 'To the things themselves!' The ambition expressed in this formula would be anathema to Kantians. Kant had, in Husserl's view, introduced philosophy to an adequately-conceived concept of the a priori. But in doing so he had placed the phenomena with which philosophy must deal beyond the limits of direct human enquiry. Husserl's ambition then was to find a way to make accessible what Kant had declared inaccessible, to make the fundamental structures of human consciousness objects of *experience*, rather than of argument and inference alone.

Husserl's second major criticism of Kant is related to this first. Kant's transcendental deduction of the categories had suggested the possibility of grasping the fundamental constitutive structures of human consciousness. At the same time however it had tended to deflect attention from the task of uncovering the sorts of ideal objects that could potentially be grasped in a direct intuition. Thus again Kant's philosophy seemed to be giving with one hand and taking away with the other. It is one thing to show that certain pure concepts are presupposed in all experience. But what of the non-categorial elements that had historically been the focus of philosophical attention? How was Kant's method going to help us isolate non-categorial essences – disputes concerning which lay at the basis of so many controversial philosophical questions? Kant had suggested the possibility of a systematic enquiry into the a priori components of our knowledge, but had then apparently closed the door to the type of enquiry that would yield genuinely useful results.

Husserl's criticisms of Descartes take a different form. Husserl credits Descartes with the discovery of what he terms the 'transcendental ego' – the self that is held to play a constitutive role in respect of experience. Having made this discovery however, Descartes, in Husserl's view, immediately misinterprets it. In his Second Meditation Descartes stands on the verge of an exploration of the structures of the transcendental self. But he makes the mistake of treating this self as if it were a '*tag-end of the world*'.[13] Descartes displays none of the pessimism that Kant was later to show concerning the prospects for a systematic exploration of the structures of the transcendental ego. But he is unable himself to carry this exploration through, because he fails to appreciate the nature of his own discovery.

In order to avoid the errors and shortcomings of his predecessors, Husserl determines to pursue what he refers to as the *criticism of transcendental experience*.[14] This will amount to an exploration and investigation of the essential structures of human consciousness.[15] But the field of experience to be investigated will need to be suitably prepared. In order to allow consciousness to yield up its secrets in the right kind of way, Husserl argues, it must be 'reduced'. That is, we must prepare our experience in such a way that we can view it as it is in itself, rather than considering it with all of its empirical entanglements. The proper approach will be to consider experience with respect to its *sense*, while 'bracketing', or 'putting out of play' any implicit commitment to the actual existence of any of the objects of experience.[16]

---

[13]Husserl 1995, p. 24.

[14]Ibid., p. 29.

[15]Ibid., Second Meditation.

[16]Ibid., section 8. See also Husserl 1982, pp. 57–62.

This phenomenological bracketing of ontological commitments, or *epoché*, sounds complex, but in essence it is not. The nature of the *epoché* is best explained in relation to sense experience, and by thinking of the senses (as Descartes in fact seems to have conceived of them) as providing us with sensory messages or 'reports'. Suppose I am walking through a wood. The experience is conveyed through a rich stream of sensory messages. The sunlight is projecting a complex pattern of light and shade through trembling branches. Leaves caught in shafts of light glow in a variety of vivid greens. There are deep shadows between the trunks, stretching into the distance and out of sight. The ferns that cover the ground are damp, and beaded with dew. The leaf mould beneath them is filling the air with the scent of decay, and yielding beneath my feet as I walk. Every element in this complex of impressions can be thought of as a message conveyed to the mind by the senses. For example, 'the leaves are vivid green'; 'the leaf mould underfoot is yielding'. In order to focus on what is directly present to consciousness here, Husserl would argue, it is necessary to reduce these messages to something 'meant' in consciousness.[17] The sensory report that the leaves are vivid green could conceivably be false. For example, some strange trick of the light could be at work making them appear green when in fact they are yellow. But it is nonetheless impossible to doubt the *content* of the report. My senses are undoubtedly telling me that the leaves are green, whatever colour they may be in truth. It is therefore possible to grasp the sense of the report with total confidence, while remaining completely uncommitted on the question of its truth. It isn't necessary to believe I actually am brushing through damp ferns to grasp the content of the experience. Indeed, the whole experience can be re-interpreted in such a way as to reveal its content, while bracketing all ontological commitments: A complex pattern of light and shade through trembling branches. Vivid green leaves caught in the light. Dark shadows stretching beneath the trees. Ferns brushing damply against my legs. Leaf mould yielding beneath my feet. All this is undoubtedly present to my consciousness, and makes up the content of my experience, whatever the truth may be concerning what is actually happening.

Each of these 'reduced' sensory messages has one or more objects. For example, the message that refers to ferns brushing against my legs has ferns as its object. But since I might conceivably have an experience with an identical content even when I am in fact lying in bed and dreaming, we cannot properly think of these objects as being identical with the physical objects that we usually regard as causing such experiences in us. The 'object' in the present sense is ultimately inseparable from the sensory experience, and the experience refers to this object purely in virtue of having the particular content it has – in much the same way as the non-existent characters of a work of fiction are ultimately inseparable from the work in which they are described, and are referred to purely and simply in virtue of the content of the work. (By contrast, a historical novel might succeed in referring to a real historical figure despite the fact that, as regards its content, the novel misrepresents that figure in a variety of ways.) There is, for Husserl, no necessary link between

---

[17]Husserl 1995, p. 33.

the experience, with these its 'intentional' object(s), and any particular object in the physical world. Indeed, there is no necessary link with the physical world at all. We cannot, for example, infer the existence of a physical world from such an experience.

Experience thus reduced to its content, or sense, constitutes the field in which Husserl wished to conduct his explorations into the fundamental structures of consciousness. Every such experience will have its specific intentional object(s) and by various further operations, including varying the intentional object in imagination, Husserl considered it possible to identify what would amount to a fundamental conceptual 'vocabulary' of possible human experience.[18] By focusing on and exploring such experiences, he argued, we can identify essential structures of human experience – 'essences' – which delimit the range of possible experiences that are open to us.[19]

Husserl held that while every act of consciousness is oriented toward a corresponding intentional object, there is no fundamental division between the intentional object and the act that 'intends' it. Rather, the act is thought of as effectively *projecting* an intentional object, as its 'meant' content. Ultimately, the intentional object is simply the effect of an intentional act.[20] Thus the duality of act and intentional object is more apparent than real, because the object is tied inseparably to the act of consciousness (or, more precisely, to acts of consciousness of a particular type). It follows that the systematic exploration of the range of possible intentional acts will ultimately tell us all we need to know about the possible contents of consciousness.

Although Husserlian phenomenology is not officially solipsistic, its general orientation evidently tends toward solipsism. Unsurprisingly then, the problem of providing a subjective foundation – a foundation in consciousness – for our knowledge of an external world, presented Husserl with grave difficulties.[21] The source of this problem is the doctrine that the intentional object is nothing but the intentional effect of the corresponding intentional act. Evidently, when the act is not one of simply perceiving an object, or entertaining an idea, and involves some kind of affirmation, the straightforwardly conceptual content of the act has to be supplemented by categorial elements (for example, a concept of things *being* thus and so), which seem not to belong straightforwardly to the intentional content of the act. Consider for example the two acts of seeing a red door, and judging that the door is red. These would seem to share the same intentional object – the red door. But the two acts indisputably differ in their meant content. This difference is revealed if we try to substitute the proposition 'the door is red' for the noun phrase 'the red

---

[18]Ibid., p. 74.

[19]Husserl 1964, p. 6. Husserl's methodological innovations were, I should emphasise, far more elaborate than the above sketch can hope to convey – though a sketch is sufficient for present purposes.

[20]Husserl 1995, section 20. See also p. 65: 'Each object that the ego ever means, thinks of, values, deals with, likewise each that he ever phantasies or can phantasy, indicates its correlative 'system', and exists only as the correlate of its system'.

[21]Ibid., sections 25–8, 58–62.

door' in a typical sentence in which the latter might occur.[22] 'The red door is open' is grammatically quite acceptable, but 'the door is red is open' is not. The difference in sense between the noun phrase, and the related perceptual act, and the affirmation that the door is red, and the related act of judgement, is palpable. And yet it seems not to show up in any difference in the relevant intentional objects. It seems then that the intentional act of affirming that the door is red somehow manages to incorporate a categorial element – a sense of the door's *being* red – which is indispensable to its sense, but is no part of its straightforward intentional content. Husserl concludes that it is necessary that our sensory 'intuitions' (that is, our sense perceptions), should include a categorial element.[23] To fulfil an expectation of a door's being red, and thus verify the affirmation that the door is red, I must be able not only to perceive or 'intuit' (e.g.) red doors, but to intuit a door's *being* red.

In his early *Logical Investigations* Husserl discusses the problem of categorial intuitions at length, without developing a satisfactory solution to it.[24] Husserl's discussion of this topic captivated the young Heidegger, who was later to become Husserl's assistant, and ultimately his most influential successor. Heidegger was aware, from his reading of Aristotle, and from the writings of Husserl's teacher Brentano, of serious potential difficulties with the idea of an intuition of *being* as such.[25] Husserl's general approach (even at this early stage) seemed to imply that *being*, in all of its many senses, ought to be reducible to something 'meant' in a corresponding act of consciousness. But at the same time, the puzzle of categorial intuitions is that this cannot apparently be done, since:

> The form giving flexion 'Being', whether in its attributive or predicative function, is not fulfilled. . . in any percept.[26]

The moral that Heidegger drew from Husserl's puzzle of categorial intuitions, and its apparent relation to traditional Aristotelian concerns with the manifold senses of *being*, was that the problem of 'the meaning of *being*' had the potential to shatter the Husserlian conception of consciousness as, ultimately, a self-contained unity. *Being*, and the problem of its 'meaning', seems necessarily to refer us to an 'outside', to something that transcends consciousness.[27] Whether or not this outside is conceived of as a physical external world, it seems that the object to which any understanding of *being* relates cannot be a mere projection or correlate of an act of consciousness.[28] Given this, the Husserlian vision of phenomenology as proceeding to a systematic exploration of essential structures of human consciousness on the basis of the methodological assumption that the nature of the intentional object

---

[22]Husserl 1970b, pp. 632–3.

[23]Ibid., pp. 780–2.

[24]Ibid., Investigation 6.

[25]Not the least of these is that *being* seems (as noted by both Aristotle and Brentano), to have a number of different senses, which are not related to each other in any obvious way.

[26]Husserl 1970b, p. 780.

[27]Heidegger 1988, sections 8 & 9. See also Heidegger 1962, section 43.

[28]Heidegger 1988, section 9c.

can be satisfactorily analysed as, in effect, a projection of an intentional act, seems like an unrealisable dream. There is more to human knowledge and consciousness than a potentially intra-mental relation between intentional objects and intentional acts. Consciousness must be conceived of as *essentially* self-transcendent.[29]

Accordingly, in his own magnum opus *Being and Time*, Heidegger suggests that the real scandal of the repeated attempts philosophers have made to prove the existence of an external world is that it mistakes the nature of the philosophical task. It is not that the evident self-transcendence of consciousness is sufficient to show that an external physical world must exist. Rather, the point is that epistemological doubts concerning an external world are otiose, given the implications of the self-transcendence of consciousness.[30] The 'objects' with which consciousness deals cannot coherently be thought of as mere projections of psychic acts. The intentional object of a thought concerning the railway station at Marburg, for example, is not some psychic phantom, but the particular railway station in question itself.[31] Acts of consciousness may be essentially 'intentional', but their intentional objects typically transcend consciousness, and must do so if we are to experience them as *being* thus and so.

The Heideggerian conception of human consciousness is therefore quite different from that of Husserl. For Husserl, the apparent duality between the intentional object and the experiencing subject cloaks a fundamental unity. They are ultimately nothing but subject and object 'poles' of a series of intentional acts.[32] For Heidegger, the intentional object is precisely not something that can be reduced to an intentional act. Instead, consciousness has to be conceived as *ek-static*, literally 'beside itself', 'escaping itself' into the world.[33] Heidegger does not altogether abandon Husserl's project of a systematic analysis of structures of human consciousness. But in the light of his doctrine of the self-transcendence of consciousness, the task is reconceived. Human consciousness typically transcends itself toward some 'worldly' object. Any fundamental structures it conceals cannot therefore be studied in a transcendentally reduced consciousness, but must be evident instead in the various intentional relations that obtain between ourselves and the worldly objects with which we interact.[34] Thus the key philosophical task, as Heidegger conceives it, becomes that of conducting an 'existential analytic', a systematic analysis of the essential structures of (what we should probably no longer refer to as) 'consciousness', as revealed in the everyday intentional acts that go to make up our worldly existence.[35]

---

[29]Heidegger 1992, sections 11a, 11b.

[30]Heidegger 1962, section 43a.

[31]Heidegger 1988, p. 70.

[32]Husserl 1995, section 20.

[33]Heidegger 1988, p. 267.

[34]Correspondingly, the traditional metaphysical distinction between essence and existence is undermined, at least insofar as it rests on the claim that essence and existence are fundamentally distinct.

[35]Heidegger 1962, part 1.

The particular focus of Heidegger's existential analytic is the problem that proved fatal to the Husserlian approach: that of elucidating the meaning of *being*. In light of the Heideggerian re-conception of the nature of the task however, the entire project takes on a somewhat paradoxical air. The fundamental structures of human 'consciousness', and in particular the structures that pertain to the question of the meaning of *being*, are to be explored by means of an existential analytic. And yet 'existence' is itself one of the many meanings of *being*. Doesn't the entire enquiry then presuppose access to something called human 'existence', which could only really be properly accessible on the basis of the understanding of *being* that it is supposed to issue in? Heidegger acknowledges that there may appear to be a problem here, but he re-interprets the apparent problem as a virtue of his preferred approach. It is distinctive of a genuinely human approach to thinking about *being* that our own existence is, and will remain, 'an issue' for us.[36] We cannot pretend we will ever be done with the matter of thinking through, reflecting on and questioning, the nature of human existence. At the same time, if we are ever to have even provisional answers to the problem whose solution eluded Husserl (that is, the problem of categorial intuitions, and in particular of the meaning of *being*), we will have to seek those answers not in a phenomenologically reduced consciousness, but in human life as it is actually lived. The problem sounds circular, but the circle involved is not a vicious one, according to Heidegger, because we are seeking answers to questions whose solutions we, in a sense, already have.[37] The task is not that of solving a riddle concerning a mysterious and arcane topic. Given the reflective nature of human beings, we are posing a question whose answer could hardly be closer to us. The existential analytic must therefore be conceived as proceeding hermeneutically. The interpretive nature of the enquiry is such that it presupposes that we already have some sort of answer to the question we are setting out to address. But in the particular case in question, if nowhere else, this is surely something we are entitled to assume. One would have to have a very strange idea of both human existence, and the question of the meaning of *being*, to think that the answer to that question was not one with which we are already intimately, even if still obscurely, familiar.

Partly in order to acknowledge the paradoxicality of the task of analysing the 'existence' of a being whose distinguishing characteristic is that its own *being* is permanently in question, Heidegger refers not to 'human being' – since that term would imply some (perhaps minimally) taken-for-granted conception of human nature – but to *Dasein*, literally 'there-being'.[38] This is not an arbitrary terminological innovation. It aims to express the distinctive self-transcendence that provokes the enquiry itself: What is this being that always seems to find itself located in some 'there', always-already outside of itself? – not groping towards an external world, but already 'thrown' into a world, inextricably entangled with a range of objects, physical and psychic? The picture is of a being whose *being* is an issue for

---

[36] Ibid., section 9.

[37] Ibid., pp. 194–5.

[38] Ibid., p. 27.

it precisely because it cannot achieve the sort of comfortable self-contained stability that the Husserlian conception of the transcendental ego, with its intentional acts and objects, suggests. One aspect of *Dasein's* being 'thrown' into its world is that it can never be found without some understanding of *being*. This is the source of the inescapable circularity of the enterprise of existential analysis, but also of its legitimacy. How else could the task be approached? And who but a being already equipped with *some* understanding of *being* would, or ever could, initiate it?

8.3 Dasein, Intelligibility and Alienation

Heidegger's existential analytic is held to be possible only because the *being* of *Dasein* is 'being-in-the-world'.[39] By this, Heidegger means to suggest both that *Dasein's* existence has the character of being always already involved in a world of objects and concerns, and also that this sort of involvement implies some under-standing of *being*.[40] This general picture is more or less dictated by Heidegger's conception of the *ek-static* nature of *Dasein's* existence. The manner in which consciousness is related to its objects is such that perfect coincidence-with-self is unachievable for us.

Referring back to the discussion of self-actualisation and alienation in the previous chapter however, we might be forgiven for concluding at this point that Heidegger is suggesting that human existence is destined to be irremediably alien-ated.[41] Perfect self-actualisation would be a condition in which we achieved a certain self-coincidence: in which we 'became what we are', and our *being* thereby ceased to be an issue for us. Evidently, Heidegger thinks this sort of condition is unachievable, by us. But to conclude that, on Heidegger's view, alienation must be inevitable, would be to ignore the paradoxicality to which his conception of his enquiry as a hermeneutic one, and his use of the term *Dasein* in place of 'human being', is intended to point. The reason why we cannot declare that Heidegger's conception of human existence is one on which we are condemned to self-alienation is that, in his model, there is nothing that corresponds to Hegelian self-certainty or 'self-actualisation'. Heidegger rejects the idea that there could possibly be any ulti-mate human coincidence-with-self. Thus he implicitly rejects the idea of human self-actualisation, as envisaged by Hegel, amongst others. Insofar as Heidegger can be said to hold that human beings have an essence, he effectively thinks we are *essentially* alienated.

The account of being-in-the-world – the state of 'being outside of ourselves', thrown into a world, in a manner that ensures that our own *being* never ceases to be an issue for us – is as close as Heidegger ever comes to defining what it is to be

---

[39] Ibid., pp. 33–4.

[40] Ibid., section 12.

[41] See above Section 7.1, and Cooper 1990, pp. 34–6.

human. But it is, in a way, senseless to talk of this as a condition of 'alienation', because this just is what it is to be human, on his view. On the other hand, this state can hardly be characterised as one of self-actualisation, since its most characteristic quality is a lack of any stable self-certainty or self-possession. If by 'alienation' is meant some state of being 'beside oneself', unable to achieve the comfortable coincidence-with-self that the essentialist model of the overcoming of alienation suggests, then it seems we must say that Heidegger thinks that we are condemned to alienation. But although Heidegger's philosophy is to some extent a philosophy of resignation, it does not follow that he thinks we should simply resign ourselves to such apparent alienation. Rather, the situation seems to call for a more adequate conception of alienation. Heidegger's philosophy denies us the sort of resolution that is promised by essentialist conceptions of alienation. But for all that it does not render the idea of an overcoming of alienation entirely meaningless. There is after all a difference between saying that we are condemned to a state in which our *being* remains an issue for us, in which we can never achieve perfect coincidence with any supposed human essence, and saying that we are condemned to live with a damaged or distorted self-conception.

According to Heidegger, *Dasein* lives 'in the truth' with respect to its world.[42] That is to say, it lives in a world populated not simply by objects of experience, but by objects of knowledge and understanding. A relevant question then is: *how* does *Dasein* come to find itself in an intelligible, knowable world? This question can be addressed by considering a contrast case. Why is it that the world sometimes loses its accustomed intelligibility? In a state of acute anxiety the world of our experience is drained of its usual significance. We cannot settle to the things that usually interest us, we look at them blankly, and they seem to reflect this blankness back to us.[43] Heidegger's interest in this unusual state does not seem to be motivated by any conviction on his part of the independent value of a phenomenological analysis of anxiety. Instead, his hope is that the analysis of anxiety, a state in which the world seems suddenly meaningless, will shed light on the concernful dealings with worldly things that are more typical of our everyday being-in-the-world. In particular, he hopes that it will shed light on the origins of the (normally) taken-for-granted intelligibility that characterises such everyday being-in-the-world.[44] Such everyday intelligibility is clearly revealed in the context of manual labour. All manual work involves complex structures of assignments and instrumentalities. It is on the basis of these structures that the individual elements within the complex first show themselves *as* that which they are.[45] For example, a hammer first appears *as* a hammer in the context of the workshop, and on the basis of the uses the hammer is put to there. It is especially when we are working with it that the hammer is present

---

[42]Heidegger 1962, p. 270. See also Heidegger 1988, p. 18.

[43]Heidegger 1962, section 40.

[44]Ibid., sections 39–41.

[45]Ibid., section 15. The immediacy of the kind of understanding we have of a tool in actual use is vividly captured in the *Odyssey* book 19 'Iron itself can draw men's hands' (Homer 1965).

for us *as* a hammer, in connection with the particular practical project on which we are engaged.[46] But if something disrupts the work, and the associated complex of assignments and goals – for example if the hammer breaks – this intelligibility evaporates. The broken hammer just lies there as a dumb thing – it no longer announces itself in experience as a hammer.[47] It ceases to 'speak' to us of anything.

Heidegger takes a similar approach to the intelligibility of linguistic signs and tokens. Their intelligibility, and the type-token relations they display, are founded on a similar instrumentality.[48] And a broadly similar analysis is provided of the foundations of our understanding of the natural world, and natural resources.[49] Finally, a not-dissimilar analysis is applied to our understanding of others. Heidegger recognises that the basis of our understanding of others differs from that of our understanding of equipment and signs. Our encounters with others are always pervaded by a sense of being in the presence of someone who is another-like-me.[50] Nevertheless, our understanding of them is also conditioned by instrumental factors, to the point where we frequently understand them in terms of their function – 'they are what they do'.[51]

Tracing the intelligibility of *Dasein's* everyday world to the context of use is central to Heidegger's existential analytic. At the same time however, the finding that 'essences' are primarily evident in the context of use suggests a dual danger. Firstly, it suggests that we will always tend to interpret the beings with which we interact primarily on the basis of their various shifting functions as objects of use.[52] Secondly, it suggests we will tend to be complacent about the ways in which we interpret them. The second of these dangers follows from the first, by a process that we should now consider.

That a hammer is able to appear to us as a hammer is a function of the way in which the workshop, with its tools and materials, is arranged, with a view to the completion of some task we have set ourselves. Thus the hammer's *being* a hammer is not a truth independent of some instrumental arrangement of tools and materials, with a view to some practical end.[53] To this extent, the *being* of the hammer is closely bound up with our own *being*, since the projects in relation to which the

---

[46]Heidegger 1962, pp. 98, 189–90. See also Mulhall 1990, chapter 4.

[47]Heidegger 1962, section 16.

[48]Ibid., section 17.

[49]Ibid., p. 100.

[50]Ibid., p. 154.

[51]Heidegger 1962, pp. 153, 163: '[W]ith the equipment to be found when one is at work, those Others for whom the work is destined are "encountered too". If this is ready-to-hand, then there lies in the kind of Being which belongs to it (that is, in its involvement) an essential reference to possible wearers, for instance, for whom it should be "cut to the figure"'; 'In that with which we concern ourselves environmentally, the Others are encountered as what they are; they are what they do'. See also Heidegger 1988, pp. 289–90.

[52]See Heidegger 1977b. The pervasive modern interpretation of nature as nothing but a resource is a key theme of this later essay.

[53]Heidegger 1962, section 15.

hammer is a device for hammering do not exist independently of human beings. Nevertheless, it is possible to lose sight of this relationship between the *being* of the beings with which we have everyday commerce, and the human projects which are ultimately responsible for that *being* (indeed in our typical state of life we do not, according to Heidegger, have this dependency clearly in view).

Where it is only a matter of overlooking the human origins of the *being* of some tool, or other item of equipment, the consequences of this habitual myopia are relatively mild. Nevertheless, the *being* of the object, revealed through use, will be potentially at odds with other ways of understanding it. For example, the natural world that the peasant views as a resource will appear different from the natural world that is studied by the scientist.[54] This may not be a problem in traditional forms of life, but if the peasant's way of looking at nature is harnessed to modern industrial technologies, then a dangerous view of the world as universally a resource may potentially arise from it.[55] The general moral is that although typical ways of understanding beings may derive from instrumental attitudes, and although this may not be a problem in many contexts, it will be likely to become a problem under rapidly changing conditions. Instrumentality as such is not necessarily a problem, but complacency is – and there is a danger of complacency in all such situations because we are always apparently able to 'read off', from a given instrumental complex, the *being* of the various elements that compose it.[56]

The danger here is redoubled where humans themselves form a part of the relevant instrumental complex, for here we make a return journey from a certain understanding of the independent origin of the *being* of the objects with which we have dealings, to a related understanding of our own *being*. Once we are in the habit of 'reading off' the *being* of the objects with which we have dealings from those objects, as if it belonged to them independently of ourselves and our own projects, it becomes inviting to read off our own *being* from them.[57] Thus a shoemaker who grows so used to his workshop environment that he treats his tools as if they are what they are, independent of the uses that humans have devised for them, is likely to interpret himself along similar lines. That he is a shoemaker will appear to be a truth as self-evident and as natural as the truth that his hammer is a hammer. And yet, while there is evidently an everyday sense in which it is true to say of the shoemaker that he is a shoemaker, such a judgement risks a form of self-stereotyping. If it is humans who are ultimately responsible for the *being* of shoes, hammers etc., nobody is a shoemaker from fate, from nature, or from destiny, but only in relation to some set of human projects. To naturalise the *being* of the equipment with which one deals is one kind of error. But to go on, on the basis of this error, to naturalise

---

[54]Ibid., p. 100.

[55]See Heidegger 1977b.

[56]Heidegger 1962, section 15. The arrangement of the workshop, or a constellation of resources, comes to imply a form of life.

[57]See Heidegger's discussion of 'Inauthentic self-understanding by way of things' in Heidegger 1988, pp. 289–90.

one's own *being*, by reading conclusions about one's own nature back from the equipment with which one is surrounded (for example, by inferring, on the basis of the *being* of the objects with which one deals, the nature of one's proper interactions with them), involves a more pernicious type of error. It is also, evidently, an error which we can make with respect to others, as well as ourselves. When we come to consider the *being* of others, they may be referred to the instrumental complexes within which we encounter them.

All of these errors, in which the *being* of an object or a person is naturalised, as if it were something free-standing, something independent of human choices and human projects, involve a malaise that Heidegger terms 'inauthenticity'.[58] Heidegger denies that inauthenticity (and its logical complement, authenticity) are ethical concepts.[59] Nevertheless it is clear that inauthenticity is at the very least a proto-ethical concept, being recognisably a form of alienation. And yet it is alienation of a strange kind. Referring back to our earlier discussion, it is evident that while Heidegger can hardly be called an essentialist, he does take the view that a distinctively human existence is *ek-static*: 'self-actualisation', as ultimate coincidence-with-self, simply is not possible for us (nor, presumably, even desirable). And yet if human existence can be said, even loosely, to be *essentially ek-static*, there remains a sense in which self-actualisation is achievable for us, so long as the *ek-static* character of our existence is preserved. If resoluteness in the face of our *ek-static* existence is the Heideggerian counterpart of self-actualisation, as Division Two of *Being and Time* suggests, then inauthenticity becomes, in effect, the Heideggerian proxy for alienation.[60]

A considerable reversal has taken place here then. Everything, from Heidegger's departure from Husserl, and the introduction of the idea of the self-transcendence of consciousness, has been pointing in the direction of the conclusion that a certain sort of candour with respect to ourselves, and the inescapability of our own *ek-static* nature, is the proper accompaniment to our self-understanding. On this sort of view, self-actualisation as traditionally conceived, as a kind of comfortable self-possession and self-coincidence, is an escapist fantasy. This implies that certain monistic conceptions of self-actualisation express the ultimate in human alienation, taking what in practice would amount to an alienated condition as an ethical ideal.[61] The distinctive terminology of authenticity and inauthenticity is necessary then to highlight the reversal that has taken place. What Heidegger refers to as 'authenticity' would be more or less indistinguishable from alienation, on some views of alienation; while what he terms 'inauthenticity' is apt to look like Hegelian self-actualisation, at least in its more mundane forms. In a sense then, Heidegger is

---

[58] Ibid. See also Heidegger 1962, section 27.

[59] See Heidegger 1977a, p. 212. Sartre however considered Heidegger's position on this to be disingenuous. See Sartre 1989, p. 80.

[60] Heidegger 1962, sections 54–60.

[61] Such at least is what an existentialist reading of Heidegger would tend to make of, for example, traditional Aristotelian conceptions of human flourishing, and Hegel's conception of self-actualisation in the ethical state.

reversing a more traditional conception of self-certainty and alienation. And yet, there remains an undeniable quality of self-possession attaching to Heideggerian resoluteness in the face of the way we find ourselves pitched into a world that is not of our choosing or making. Authenticity seems above all to consist in *facing up to* our (apparently) *essentially* alienated condition. And while this makes it something very different from Hegelian self-certainty, it clearly amounts to a victory over some familiar forms of alienation – such as the various forms of complacency already discussed.

Before concluding this section of the discussion it is important to mention a couple of points of clarification. Firstly, while it may sound as if our understanding of the *being* of the beings with which we have dealings, whether human or not, is a very pragmatically-oriented affair – essentially related to the context of use – Heidegger does not in fact take such a narrow view. It happens that the context of the workshop, and the use of tools, serves as a particularly clear illustration of the way that the *being* of the objects with which we have dealings derives ultimately from particular human practical projects. But our interaction with our world is never wholly pragmatic, and Heidegger nowhere suggests that it must be interpreted in a narrowly pragmatic fashion. Secondly, while it was convenient to couch the above account in terms of what we 'take', 'understand', or otherwise interpret beings to be, it should be borne in mind that when Heidegger is talking of our interpretation of the *being* of beings he is not necessarily, or even primarily, thinking of what we explicitly believe them to be. Rather, he is thinking of the fundamental pre-theoretical relations we have with them, that make it possible for us to have explicit beliefs about them in the first place. If I did not have a pre-theoretical grasp of what it is for something to be a hammer, for example, I could not arrive at the belief (whether correct or incorrect) that a given implement is a hammer. At issue here is not the existence and nature of particular explicit interpretations of the world, but the existence and nature of the basic structures of intelligibility that serve as a foundation for more elaborate systems of belief.[62]

8.4 Inauthenticity and Objectification

Heidegger's analysis of inauthenticity has clear relevance to our earlier analysis of ideological stereotyping. The inauthentic self-understanding of the individual who 'reads off' his own *being* from the workaday world of his everyday existence combines a self-instrumentalisation with a debased and debasing form of self-understanding. It amounts to a self-*objectification*, as that term was interpreted in Chapter 5. At the same time, inauthenticity is not restricted to this affair of self-interpretation. It is perfectly possible to behave inauthentically in one's interactions with others, drawing one's understanding of them from the instrumental roles they

---

[62]Heidegger 1962, p. 90.

are called upon to perform.[63] The inauthentic human individual, who draws her self-understanding from the instrumental complexes with which she is surrounded, is 'lost' in her world.[64] But this is not a world that is devoid of reference to other human individuals. *Dasein* does not simply exist in a world of instruments, which happen to be available to others. It belongs to the very meaning of these instruments that they are available to others. Heidegger refers to this aspect of being-in-the-world as *mit da-sein*, or *mitsein*, 'being-with'. It is not an accidental fact about *Dasein* that it happens to find itself in a world populated by others who are encountered as 'another like me'. Rather, it belongs to *Dasein's* 'essential' constitution that this is so – this fact colours *Dasein's* relation to its world, since it belongs to the sense of this world that it is a public world, and that the instrumental items it contains are available for use by others.[65] What may at first look like a rather individual-istic and even solipsistic model of inauthentic being-in-the-world is destined then to become something quite different, once this reference to others is noted. It isn't the individual human being who gives meaning to the instrumental complexes she encounters. Rather, they refer essentially to human social activities and projects. This element of availability to, and thus reference to, others – 'being-with' – does not however automatically lead in the direction of adequate self-understanding and authenticity. Indeed, the reverse is typically the case. That the everyday objects of use that fill our lives are essentially available to others compounds our tendency to inauthenticity, since each of them implies a 'they' to whom they are available, and whose activities give them meaning. Surrounded by an amorphous 'they', to whom the instrumental complexes with which we deal also refer, the abdication of responsibility for our own *being* becomes much easier. One conducts oneself in all things as 'they' do, and thus the presence of the 'they' functions as a further fac-tor leading to inauthenticity.[66] Heidegger thus apparently takes the view that in any context in which inauthenticity threatens, the essentially social dimension of *Dasein* represents a route to an intensified inauthenticity. Not only is some kind of 'self-instrumentalisation' a constant danger, but, surrounded by a world that presents us with a constant running commentary concerning what 'one' does, and how 'one' thinks, *Dasein's* essentially social existence further facilitates inauthenticity.

Heidegger says remarkably little in all of this concerning the effect that an indi-vidual's self-understanding, as mediated by the 'they-self', has on *others*. As pointed out above, he does refer briefly to the fact that inauthentic *Dasein* tends to interpret others according to their role.[67] And this would be quite consistent with the idea that inauthentic *Dasein* interprets itself according to its (now social) role. This remark

---

[63]To be clear: the claim here is that objectification, as previously discussed, is a form of inau-thenticity. I do not mean to suggest that all forms of inauthenticity can be analysed as forms of objectification.

[64]Heidegger 1962, p. 107, and passim.

[65]Ibid., sections 26, 27.

[66]Ibid., section 27.

[67]Ibid., p. 163.

looks promising for our interpretation of objectification, and although Heidegger does little to develop it, it seems more than sufficient to warrant our attributing to him the view that inauthentic *Dasein* interprets (and presumably treats) others instrumentally, in interpreting itself instrumentally. It is significant that in this type of case a *partial* instrumentalisation of the other becomes in effect a *total* instrumentalisation, more or less by default. We have seen how in the traditional ethic of respect for persons the distinction between treating another as a means, and treating her as a mere means takes on considerable importance. We have also seen that in certain circumstances this distinction becomes very difficult to draw. Now, in respect of the kind of instrumentalisation of others that would be involved where an individual (inauthentically) based her interpretation of another on the manner in which that other was encountered in some instrumental context, the line between treating the other as a means, and treating her as a mere means, seems largely to vanish. On the traditional understanding of the instrumentalisation of the other I can, for example, plead innocence of reducing a café waiter to the status of a mere means if I can show that in my treatment of her I acknowledge her status as an end also, for example by tipping her generously. As previously discussed (Chapter 5), payment for services is often taken to be sufficient to suggest the existence of a consensual, non-coercive relationship, founded on mutual respect.[68] But if I behave in such a way as to draw my understanding of a human individual from her function – the manner that Heidegger views as inauthentic – then it seems useless to protest that by (for example) tipping her I recognise her as something more than a mere means. The important question in such a case surely is: given that she is to some degree treated as an instrument, how, nevertheless, is she *understood*? If the answer is that she is also understood as a means, then this seems to trump any potential appeal to contracts, payment etc. If, and insofar as, we *regard* another as a mere tool or resource, our behaviour towards her looks inauthentic, and we seem guilty of objectifying her, whether or not it is also possible to point to other aspects of the situation that suggest that she is *treated* as more than a mere tool or resource. Thus the (at the best of times vexed) distinction between reduction to a means and reduction to a mere means diminishes in significance once the question of my *conception* of the other (as opposed simply to my treatment of her) is firmly in the frame. To treat another as an instrument while at the same time drawing my understanding of her from the instrumental context in which she is encountered is to objectify her as surely as if I had reduced her to the status of a slave. What matters, both from the point of view of the analysis of objectification, and from that of the analysis of inauthenticity, is the instrumental stereotyping of human beings. Whether the instrumentalisation in question can reasonably be said to be total, or only partial, all things considered, is beside the main point.[69]

---

[68]Though not necessarily in all cases, for example prostitution.

[69]None of which should be taken to imply that there do not exist even more powerful ethical reasons for avoiding the total instrumentalisation of human beings than those associated with the avoidance of objectification. To point out that objectification may be in evidence even in cases in which the victim can be shown not to have been *treated* as a mere means is not to deny that

In addition to lending itself to the interpretation of objectification I have been concerned to develop, Heidegger's analysis of inauthenticity shows how the analysis of objectification can be detached from the entrenched debates that surround straightforwardly essentialist concepts of alienation and self-actualisation, and liberal misgivings about contestable moral ideals and conceptions of human nature. It is quite unnecessary to espouse any conventional essentialist ideas about human nature to recognise Heideggerian inauthenticity as a form of instrumental stereotyping of oneself and others. The failure of an instrumentally-derived conception of another to reflect what they are is not a failure whose analysis presupposes access to a positive teleological conception of human nature. At Heidegger's hands, it presupposes only that to be a human being is to be the kind of being whose nature can never be captured by any instrumentally-derived understanding. To be human is to be the type of being whose *being* is always at issue. It is the reduction of another to some (apparent) fixed and abiding (instrumentally-derived) essence which constitutes the stereotyping element in this case, not the matter of conceiving the other along lines that fail to do justice to their underlying 'human nature'.

It is of course open to the anti-essentialist to object that Heidegger's concept of inauthenticity is itself an essentialistic concept, which implies that it is essential to human beings that our *being* should always be an issue for us. But such 'negative essentialism' is not at all equivalent to Hegelian essentialism. It is perhaps best characterised as a *sceptical* essentialism: just as knowledge remains an issue for the sceptic, even as she expresses doubt about our ever obtaining it, so the human essence remains, and should remain, an issue for *Dasein*, on Heidegger's view.[70] This idea that (what is in effect) the human essence will and should remain an issue for us informs his conception of authenticity, even while he denies that self-coincidence and self-possession (self-actualisation in a Hegelian sense) are possible for us.

For all the progress that this represents for our project of setting out the basis for a revised ethic of respect for persons which acknowledges the significance of ideological stereotyping, some important elements are still notably lacking. We still lack any clear sense of the social, reciprocal consequences of inauthenticity and alienation. The difficulty here is that while concerns about non-objectification, and the forms of respect properly due to persons, seem naturally to belong to the field of social morality, concerns about inauthenticity have a distinctly personal feel. Heidegger's protests that authenticity and inauthenticity are not ethical concepts seem disingenuous. It is implausible to deny that the idea of authenticity has a normative dimension, even if we are inclined (as I am not) to question its normative power. The problem however is that the quest for authenticity (or merely to avoid inauthenticity) looks on Heidegger's account to be a very personal affair. While for

---

treating someone as a mere means may be the more serious ethical failing. My concern is that the significance of objectification tends to be overlooked. I have no intention of suggesting that it is the only, or even the most serious, moral wrong connected with the instrumentalisation of persons.

[70]As it would not for, for example, a pragmatist, who has a neat reductive story to tell about what knowledge really is, which supposedly settles the issue. See e.g. Rorty 1989.

reasons to do with projects of personal perfection we may aim to be authentic, it seems hard to identify strong reasons why those not inclined to adopt authenticity as an ideal should do so. The problem here is twofold. Firstly, we still need good reasons, of a kind that would be appropriate to discussions concerning public morality rather than private perfectionism, to avoid objectifying others. Secondly, we (ideally) need some clearer reason to think that, in being objectified, an individual is harmed or wronged. In order to address this problem, the next chapter will aim to supplement our understanding of the connection between ideological stereotyping and inauthenticity with reference to the work of a thinker strongly influenced by Heidegger: Jean-Paul Sartre.

References

Cooper, David. 1990. *Existentialism.* Oxford: Blackwell.

Hacking, Ian. 1983. *Representing and intervening.* Cambridge: Cambridge University Press.

Heidegger, Martin. 1962. *Being and time* (trans: Macquarrie, J. and Robinson, E.). Oxford: Blackwell.

Heidegger, Martin. 1977a. Letter on humanism. In *Martin Heidegger: Basic writings*, ed. D.F. Krell, 193–242. San Francisco, CA: Harper.

Heidegger, Martin. 1977b. The question concerning technology. In *The question concerning technology and other essays*, ed. Martin Heidegger (trans: Lovitt, W.), 3–35. New York, NY: Harper and Row.

Heidegger, Martin. 1988. *The basic problems of phenomenology* (trans: Hofstadter, A.). Bloomington: Indiana University Press.

Heidegger, Martin. 1992. *The metaphysical foundations of logic* (trans: Heim, M.). Bloomington: Indiana University Press.

Homer. 1965. *The Odyssey* (trans: Fitzgerald, R.). London: Panther.

Husserl, Edmund. 1964. *The idea of phenomenology* (trans: Alston, W.P. and Nahknikian, G.). The Hague: Martinus Nijhoff.

Husserl, Edmund. 1970a. *The crisis of European sciences and transcendental phenomenology* (trans: Carr, D.). Evanston: Northwestern University Press.

Husserl, Edmund. 1970b. *Logical investigations* (trans: Findlay, J.N.). London: Routledge and Kegan Paul.

Husserl, Edmund. 1982. *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy* First Book (trans: Kersten, F.). Dordrecht: Kluwer.

Husserl, Edmund. 1995. *Cartesian meditations* (trans: Cairns, D.). Dordrecht: Kluwer.

Kripke, Saul. 1980. *Naming and necessity.* Oxford: Blackwell.

Kuhn, Thomas. 1962. *The structure of scientific revolutions.* Chicago, IL: The University of Chicago Press.

Mulhall, Stephen. 1990. *On being in the world: Wittgenstein and Heidegger on seeing aspects.* London: Routledge.

Plato. 1954. Euthyphro. In *The last days of Socrates* (trans: Tredennick, H.). London: Penguin.

Popper, Karl. 1945. *The open society and its enemies*, Vol. 2. London: Routledge and Kegan Paul.

Popper, Karl. 1963. Three views of human knowledge. In *Conjectures and refutations*, 97–119. London: Routledge.

Putnam, Hilary. 1993. Meaning and reference. In *Meaning and reference*, ed. A.W. Moore, 150–161. Oxford: Oxford University Press.

Rorty, Richard. 1989. *Contingency, irony and solidarity*. Cambridge: Cambridge University Press.

Sartre, Jean-Paul. 1989. *Being and nothingness* (trans: Barnes, H.). London: Routledge.

Chapter 9
Sartre and Sadism

Abstract Heidegger's concept of authenticity could serve as the basis of a sceptical essentialist ethic of self-interpretation. Nevertheless, such an ethic would suffer from significant limitations. Since he rejects essentialism, it seems Heidegger must regard an authentic existence as at best meritorious, rather than in any sense obligatory. Moreover, although he is sensitive to the consequences of adopting an instrumentalised self-conception, he does not concern himself with the many ways in which we may induce such self-conceptions in others. Both Heidegger and Sartre denied that they had aimed to make contributions to ethics. Nevertheless, Sartre's accounts of 'bad faith', and the sadomasochistic features of personal relationships, provide a way to augment the normative force of a sceptical essentialist ethic of self-interpretation. Sartre shows that there is positive mendacity in inducing instrumentalised self-conceptions in others. Such mendacity is sufficiently close to the forms of rational inconsistency that underlie Kant's concept of moral obligation to serve as a focus of perfect duties of public obligation, in addition to imperfect duties of private perfection.

9.1 Objectification as a Doubly Reflexive Phenomenon

The previous chapter began to sketch how an ethic of respect for persons might be developed, along sceptical essentialist lines, to take account of ideological stereotyping and interpretive moral wrongs. The instrumental stereotyping – in other words, the objectification – of a person looks to involve a form of Heideggerian *inauthenticity*. Objectification thus understood would qualify as an interpretive moral wrong because it involves stereotyping someone 'as' such-and-such, in line with some functional role they play. The victim – oneself or another – is identified with her role in a manner that actively delimits her life-possibilities. Rather than being interpreted as, first and foremost, a unique individual, who happens to fulfil some particular functional role, she is interpreted as if she were necessarily identified with that role – as if it somehow expressed her essential nature. She is interpreted in terms of the particular instrumental complexes in which she happens to be enmeshed, rather than being interpreted in terms of her own projects and possibilities.

However, it remains true to say that Heidegger's interest in inauthentic self-understanding is as a malaise whose most significant features concern our

interpretation of ourselves. His focus is on the various ways in which an individual bases her own self-conception on some model supplied by the 'they', rather than on the ways – which are allowed for in his model, but not emphasised – in which an individual might be responsible for inducing inauthentic self-conceptions in others. Our enquiry into interpretive moral wrongs has, by contrast, highlighted the ethical importance of the various ways in which we are able (perhaps via aspects of our own self-interpretation) to detrimentally affect the self-conceptions of others.

On Heidegger's account most of us live inauthentically most of the time. Since inauthenticity involves a failure to fully develop our own personhood, if there is an ethical failure involved in inauthenticity (and as we have seen, Heidegger seems reluctant to concede that there is), it would seem to take the form of a failure to realise our higher possibilities – to develop our full potential – rather than a failure to fulfil any obligations we may have to others. Interpreting the phenomenon in this way would have the effect of diluting (though by no means altogether removing) its ethical significance. Most notably, inauthenticity as a failure to realise one's true potential could not be construed as a fully *knowing* failure. It would seem to stem more from ignorance than from (e.g.) dishonesty. As a consequence, the ideal of authenticity, if it belongs to ethics at all, seems most naturally to belong to an ethic of private perfection, rather than to one of public obligation.[1] In Kantian terms, any duty of authenticity we might be said to possess would have to be understood as an imperfect duty, less immediately pressing than the perfect duties associated with our basic obligations to others.[2] With these points in mind, it seems reasonable to conclude that tying our analysis of interpretive moral wrongs closely to the Heideggerian concept of inauthenticity would tend to restrict its apparent normative significance.

In order to bring out the full potential of a sceptical essentialist analysis of interpretive moral wrongs, we need to shift the focus slightly, and explore the idea of a *doubly* reflexive process of objectification. A doubly reflexive process of objectification would involve not simply an (inauthentic) restriction or undermining of personhood, played out solely in the agent's self-understanding, but the *knowing* restriction or undermining of personhood (one's own, or that of another): the undermining of the personhood of a victim who is at the same time implicitly acknowledged to be a self-interpreting agent. When objectification is analysed as a doubly reflexive process, it appears to have greater normative significance, because it involves a key element of mendacity. While Heideggerian inauthenticity involves a failure to develop our full potential, it cannot reasonably be presented as dishonest. On the other hand, knowingly undermining another's personhood would involve a palpable element of dishonesty.

In order to develop such an analysis it is helpful to turn from Heidegger's concept of inauthenticity to the work of Sartre, and in particular to the latter's analysis

---

[1]Discussing Heidegger's concept of authenticity, Sartre notes that such authenticity must be earned – it is not a return to an original innocence, but an individual human achievement. See Sartre 1989, p. 246.

[2]See Chapter 2, above.

of concrete relations with others. Sartre was deeply influenced by Heidegger, and in turning our attention to Sartre we will be progressing further in the sceptical essentialist direction that Heidegger opened up. At the same time though, we will be turning from the highly individualistic focus of Heidegger, to a view that places more emphasis on personal and social relationships.

Heidegger and Sartre (along with other existentialist thinkers) are united by, if nothing else, their view of the human being as that being whose own being is an issue for it.[3] To be human is, on such a view, essentially a matter of confronting, and endlessly interrogating (without the prospect of any ultimate resolution) the question of what it is to be human. The implicit paradox here threatens to become explicit in the well-known Sartrean formula according to which the existence of the human being precedes and implies its essence.[4] We are condemned to an existence in which we are nothing else but what we make of ourselves.[5]

We saw earlier that the essentialist response to instrumental stereotyping (a response congruent with both Hegel's view of the ascent to universal self-consciousness, and Marx's analysis of commodification) can either take the form of asserting that there is a simple mistake involved (the worker is not a commodity, whatever anyone may think), or the more subtle form of highlighting a fallacy of relevance (such as we saw in the cases of discrimination and stereotyping considered in Chapters 3 and 4). Here the concern would be with *ideological* stereotyping. (And here, as previously discussed, the stereotype might consist substantially of truths, but would nevertheless be distorting, in light of relevant unfulfilled historical possibilities.) Explicitly distinguishing ideological stereotyping from more straightforward forms of stereotyping enables us to contrast the complex misinterpretations of persons we find in such cases with the simpler types of misinterpretation that occur when the relevant stereotype is straightforwardly false.[6] However, the paradigmatically existentialist response to examples of suspected stereotyping would be to point out that the very idea of a *simple* misinterpretation of another human being already involves an error. No attempt to sum up the nature of a human being using, for example, the concepts of the natural and human sciences, can avoid being a distorted and distorting portrayal. To instrumentally reduce a person to some particular 'what' involves, in effect, a denial of personhood. But this is not because being a person is a matter of belonging unproblematically to some non-instrumental category (as an essentialist might have it). Rather, it is because being a person is a matter of not belonging unproblematically to any category.[7]

On this type of view, the restriction or undermining of personhood involved in objectification might as well be characterised as the illegitimate *essentialisation* of the victim: the illegitimate interpretation of the victim according to some

---

[3]Sartre 1989, p. xxxviii.

[4]Sartre 1948, pp. 26–8; 1989, p. xxxi.

[5]Sartre 1948, pp. 28, 42.

[6]See Chapters 4 and 6, above.

[7]Sartre 1948, pp. 26–8, 45.

purportedly definitive (and typically, but not necessarily, instrumentally-derived) conception of what they are. Such a response remains compatible with the idea that objectification involves ideological stereotyping. But rather than interpreting ideological stereotyping as the assimilation of an individual to a category to which they do not 'by nature' belong, it would understand all attempts to give positive (non-paradoxical) content to the idea of the 'nature' of a human being to involve a fallacy of relevance. The fallacy would consist in the fact that, though the relevant generalisations might succeed in stating certain truths, these would be true only accidentally, or as a matter of historical contingency. On this view, any purported truth about what a human being is will ultimately turn out (if true at all) to be true only in one of these two ways. In developing a concept of a doubly reflexive process of objectification I will be concerned both with factors internal to the project of satisfactorily analysing the concept, and with independent reasons for preferring an existentialist conception of the human relationship to our essence, with all its troubling paradoxicality. This chapter will also serve as a transition to what I want to present as a definitive analysis of the ethical implications of ideological stereotyping, which will be developed in Chapter 11.

9.2 Duality and Intentionality

Although it undoubtedly has an air of paradox, the Sartrean view of the human being as the being whose existence precedes its essence has a serious philosophical purpose. Sartre considers the human being to be an irreducible *duality*.[8] Probably the most familiar manifestation of the duality of the human being is the philosophical problem of the mind-body relationship. On the one hand we exist as physical objects in the world. On the other hand, we exist as consciousnesses, which are somehow (mysteriously) connected to our bodies. We are not identical to our bodies, and we are not reducible to our bodies. But neither is there any obvious part-whole relationship between mind and body. We both are, and are not, our bodies, then.[9]

The complex and puzzling relationship between mind and body implies, for Sartre, that consciousness can never be identified with, or reduced to, any empirical object or process. Consequently, it cannot legitimately be presented as entirely enmeshed in causal processes, in a manner that would limit the scope of human freedom. This view of the human being suggests an initial and fairly straightforward Sartrean concept of objectification. All forms of materialism, Sartre tells us, mistakenly aim to treat the human being as an object not essentially different from a table or a stone.[10]

But while Sartre holds that our irreducible duality rests partly on the irreducibility of the mental to the physical (his stance here reflects his oft-noted Cartesianism),

---

[8]Sartre 1989, part 2, chapter 1, section 1.

[9]Ibid., p. 326.

[10]Sartre 1948, p. 45.

he also thinks of it as occasioned, more profoundly, by an irreducibility within consciousness itself.[11] It is, according to Sartre, a hallmark of the purely physical to be self-identical – to be coincident with itself. (This self-coincidence of the purely physical is referred to by Sartre as its 'infinity'.[12]) By contrast, the human being is, on his view, never fully coincident with itself. As we have seen, to be human is to have both a physical and a mental existence; and this on Sartre's view prevents our ever being fully coincident with ourselves. But coincidence-with-self is in any case ruled out by the fact that *consciousness itself* is never self-identical.

This non-identity of consciousness stems from what is for Sartre its most unique and distinctive quality: intentionality.[13] We have already considered this notion of intentionality as the most characteristic quality of consciousness (Sections 3.2 and 8.2). Consciousness is always directed toward some mental object or content; it is always consciousness 'of' something. However, despite his evident indebtedness to Husserl's conception of intentionality as the hallmark of the mental, Sartre's view of consciousness as an irreducible duality ultimately places him closer to Heidegger than to Husserl.[14] For Sartre, consciousness is always riven between the act of being aware, and the intentional object of awareness (which may be a physical object, or may be purely ideal – e.g. a geometrical figure). Intentional objects are for Sartre (following Heidegger) more than just the object poles of psychic acts. Consequently, consciousness is always already 'beside itself': *ek-static*. Consciousness can never be fully present to itself. It can never turn up as the object of its own mental acts. Nor can it be grasped as nothing more than the subjective correlate of the intentional object. The radical gulf that separates mental acts from their objects entails that we will never succeed in fully capturing consciousness itself as an object of enquiry. The only 'mind' that could ever be an object of knowledge for us is, precisely, the mind-as-object – which is in truth only one element within the duality that is consciousness.[15]

Finally, the human being is marked by inescapable duality because it is a temporal being. Although time and history might reasonably be presented as media in which we have our being, we do not live in time and history as a fish lives in water. We are involved with our pasts, and our futures, in ways that are far more intimate than any spatial juxtaposition could ever be. We never find ourselves without a history, and we are never without plans for the future. Our past and our future partly define what we now are – for example, our enduring goals and projects can only be grasped with reference to what we have been, and what we aim to be. As temporal beings then, it can be said that we 'are' what we have been (as having-been-it). But at the same time we are never reducible to our pasts. Partly because our pasts exhibit

---

[11] Sartre 1989, pp. 76–7.

[12] Ibid., p. 76.

[13] Ibid., pp. xxxvii, 72–6.

[14] Ibid., pp. xxix–xxxi.

[15] Ibid., pp. 102–5.

this intimate connection with what we now are, through the projects we are engaged in, we always at the same time transcend our pasts, toward some envisaged future.[16]

Thus there are in Sartre's view three main facets to the human being's inability to be self–coincident in the manner of a physical object, each of which involves a form of 'transcendence': We are our bodies, but we also transcend those bodies, in our mental lives. We are our consciousness, but that consciousness is inherently self-transcending – the intentional object and the act by which we are aware of it are mutually irreducible. Finally, we are our pasts, but we also constantly transcend our pasts, toward some envisaged future. Everything that we understand as (apparently) fixed and settled in the human being – our physical bodies, our pasts, consciousness as object, Sartre refers to as our 'facticity'.[17] The various respects in which the human being exceeds its facticity go to make up its 'transcendence'.[18] As facticity, the human being partakes in being 'in-itself' – the sort of being possessed by merely physical objects.[19] But as transcendence we are also being 'for-itself'. The human being is for-itself in the sense that it pursues projects of its own, and is the bearer of associated interests. But it is also for-itself in the separate but related sense of being 'there' for itself – it is self-present, self-aware.[20] To be a human being, then, is to be more than simply in-itself, *and* more than purely for-itself. The human being is a complex of interdependent physical, psychic and temporal elements.[21]

However, for all its inescapable self-transcendence, human existence displays a fundamental unity. As previously mentioned, it is in the nature of acts of consciousness that they involve self-awareness.[22] I am aware of them as my own acts. For example, both my acts of 'reflecting' (awareness of belief), and what is thereby 'reflected' (beliefs), are, in a sense, 'me', or 'mine'. The full Sartrean picture of human existence as duality is thus of an irreducible rift traversed by an inescapable unity. On the one hand the human being is divided between the physical and the psychic. On the other hand it knows that it cannot disown either its physical or its psychic aspects. Moreover, consciousness itself straddles the rift between intentional act and intentional object – though it is at the same time aware of itself as embodying this rift. Consciousness both is and is not one. It is neither this (act) nor that (object). It is – somehow – both. But it cannot be conceived of as the simple aggregate of the two.

The fundamental duality of human existence lends plausibility to the existentialist idea that the essence of humanity is to be such that its essence is always in

---

[16] Ibid., pp. 124–9.

[17] Ibid., pp. 79–84, 116–9.

[18] Ibid., pp. 84–204.

[19] Ibid., pp. 114–20.

[20] It follows, for Sartre, that simply to be conscious is to be aware that consciousness is inescapably free – see Sartre 1989, p. 129.

[21] Ibid., pp. 120–1.

[22] See Chapter 8, above.

question; and Sartre's philosophy arises out of, and constantly orbits around, this duality. To be a human being is to be self-present – where the two terms of the duality are neither reducible to each other, nor separable from each other, and their combined irreducibility-with-inseparability are preconditions for self-presence as such. The individual human being can never fully identify itself with anything that it could take as an object of knowledge. But neither can it disown that which it knows itself to be.

As a consequence of our dual nature, and the undeniable fact that we are, inescapably, this consciousness, as well as the physical body that is accessible to empirical investigation, human existence involves a constant struggle to accept one's facticity, in all its (potentially humiliating) limitedness, while simultaneously embracing the inescapable freedom proper to consciousness. Failure in this struggle is represented by 'bad faith'.[23] To fall into bad faith is to fall victim either to the temptation to identify oneself entirely with one's facticity – one's being as an object – thereby implicitly denying one's conscious being and its freedom (or 'transcendence'); or it is to fall victim to the temptation to identify oneself exclusively with consciousness, thereby denying one's empirical limitedness.[24] Neither human freedom nor human limitedness can be denied in good faith. If this predicament lends an uncomfortable instability and edge to individual existence, its effects on social relations are also dramatic.[25]

9.3 Being-with-Others

For all the parallels between Sartre's philosophical approach and that of Heidegger, they are not without their differences. Sartre criticises Heidegger (unfairly, in my view) for suggesting that an analysis of fundamental structures of consciousness could possibly tell us something about an external world. Specifically, he criticises Heidegger for confusing the idea that 'being-with' others is fundamental to consciousness (as is evident for example in the idea of objectivity, which implies availability-in-principle-to-others) with the idea that solipsism can be shown to be false.[26] Sartre's own preferred departure point is the Cartesian *cogito* – the foundational insight that a thinker cannot possibly doubt her own thinking. Descartes, Sartre observes, shouldn't be interpreted as having attempted to *prove* his own existence, as if his existence could be anything other than contingent. The *cogito* could not (supposing we were genuinely capable of doubting our own existence) function

---

[23]Sartre 1989, part 1, chapter 2.

[24]Ibid., part 1, chapter 2, section 2.

[25]In its most dramatic forms, this predicament can give rise, according to Sartre, to an uncontrolled oscillation between hate and despair. Sartre leavens his pessimism in this regard with a hint at a possible ethics of 'deliverance and salvation' (1989, p. 412n; Conclusion, section 2), but this is a line of enquiry he notably failed to develop.

[26]Sartre 1989, pp. 244–50.

as a proof that the thinker exists – if for no other reason than that it is not inter-subjectively valid. The proper way to interpret the *cogito*, according to Sartre, is as the observation that an existent consciousness cannot but be certain of its own exis-tence, and its own mental contents.[27] Consequently, Sartre avers, the business of identifying fundamental structures of human consciousness should be approached in an essentially descriptive manner. One cannot possibly prove that consciousness must have certain features (there are echoes of Husserl's criticism of Kant here). Rather, the task is to clarify what human existence involves, in such a manner as to exclude all doubt – identifying elements that as a matter of indubitable fact condition our entire experience of the world.[28]

For Sartre however (as for Heidegger), consciousness is not a phenomenon that lends itself to being grasped purely epistemically, not even as the act-correlate of an array of intentional objects.[29] The only way to understand what consciousness is – given that, qua duality, it will never be graspable as an object – is to grasp it interpretively, from the 'inside'. One element identified by Sartre, and indeed his phenomenological predecessors, as particularly important for the type of phe-nomenological enquiry on which he is engaged, is our sense of being-with others. Philosophical interest in the problem of being-with others is stimulated in large part by the importance of the idea of intersubjective availability for such key epistemic notions as scientific objectivity.[30] Accordingly, Sartre aims to give an authoritative account of the meaning of being-with, which cannot be marshalled as an argument against solipsism, but will explain how, as a matter of indubitable fact, a sense of the 'Other' colours my experience of the world. It is this enquiry that will prove particularly fruitful for our analysis of objectification and other interpretive moral wrongs.

We often encounter others simply as objects within our world. It seems we can encounter them as mere functionaries, who do not in the least intrude upon us, and of whose independent existence, as anything other than mere objects, we are practically oblivious.[31] This sort of experience could be described as an experi-ence of other bodies as zombie-like, lacking minds. Such an encounter could not be presented as a case of encountering another self, another self-present duality, like ourselves. Alternatively, we can encounter others as bodies possessed of minds. We can observe them solving problems, doing mental arithmetic, composing poetry or playing chess. And while these examples would not satisfy the hard-line sceptic about other minds, they are satisfactory everyday examples of what we mean when we talk of encountering a 'lively mind', or engaging in a 'battle of wits' or partici-pating in a 'meeting of minds'. Thus it is possible to distinguish the way in which

---

[27] Ibid., p. 251.

[28] See for example Sartre's introduction to his discussion of the existence of others – Sartre 1989, pp. 250–2.

[29] Sartre 1989, p. xxxvii.

[30] Ibid., pp. 233, 267.

[31] Ibid., pp. 380–1.

others may be available to us simply as objects, from another way in which they may be available, as mind-body systems, as objects equipped with minds. However, neither of these cases makes the other available to me as 'another, like me'. In neither of these cases is the reality of the other-as-subject a matter of my direct experience. In watching someone solve a chess problem her mental activity is (supposing it is accessible to me at all) accessible to me as an object of study, analogous to the way in which her physical activity might be an object of study for me. But what I do not observe, what I cannot (it seems) observe, is the consciousness of the other, qua self-present duality. I cannot therefore *observe* the other as another-like-me.[32]

Nevertheless, Sartre is persuaded that a genuine sense of being-with-and-for-others is an inescapable feature of our experience, which makes solipsism untenable. He cites the experience of shame as an example in support of this view.[33] Suppose that from jealousy, curiosity, or vice, I am hunched over a keyhole, spying or eavesdropping. It is perfectly possible to be engaged in such an activity with no sense of transgression at all. I may simply be engrossed in what I am doing, and the satisfaction of whatever desire inspired me to do it. But if I am startled out of my absorption by the sound of a footstep behind my back, I will immediately be thrust into a reinterpretation of the situation. The resulting reinterpretation is likely to be accompanied by an acute sense of shame.

In a case like this, the advent of another consciousness marks an irruption that transforms my world. In particular, I become aware of an alternative perspective on my behaviour – the perspective of the other. This awareness amounts to more than just an awareness of the existence of a physically different viewpoint on my activity. Most obviously, I become aware of different possibilities for evaluating my conduct. However much satisfaction I may have been deriving from what I was doing, I will now be uncomfortably aware that others can and will evaluate my conduct differently. In connection with this, I will also be aware of different instrumental possibilities connected with the objects with which I have been engaged. For example, I may become aware of the significance of the keyhole as a means to allow or prevent access, rather than purely as a window on what lies beyond. In combination, these factors contribute to my awareness of a wholly different 'situation' (that is, as opposed to merely another perspective on the same situation).[34] I will be aware of the way that, from the perspective of the other, I and my surroundings might constitute a quite different ensemble of means, ends and potential actions, reciprocally implying one another. And yet, crucial to my awareness of this new situation will be

---

[32]Ibid., pp. 227–32. In Hegelian terms, one might say that I cannot equate merely observing the other with *recognising* the other. This is not (simply) the puzzle of the existence of 'other minds'. The point is rather that, irrespective of whether the other's mind is available to me as an object of experience, the other cannot in principle be available to me as the self-present duality she is (for this would require my being the subject of the other's experiences, and would thus involve my own annihilation). We cannot therefore coherently be claimed to derive a sense of the existence of others-like-me (the existence of other subject-object dualities) simply from observation.

[33]Ibid., pp. 221–2, 252–302.

[34]Ibid., p. 259.

the fact that I myself feature in it, for the other. I cannot of course feature in the situation for the other just as I do for myself (I am no more straightforwardly available as a self-present duality to the other than she is available as a self-present duality to me). The 'self' that is to be found in this situation – which is the situation as it is *for the other* (of which I am now uncomfortably aware) – is not the self-present duality I really am. Rather, it is an object-self:

> For me the Other is first the being for whom I am an object; that is, the being through whom I gain my objectness [. . .] If someone looks at me, I am conscious of being an object. But this consciousness can be produced only in and through the existence of the Other.[35]

Equally importantly, I cannot simply disown this object-self that I discover through the other.[36] It may not be the self that I usually take myself to be, but I cannot in good faith deny its connection with me. Crucially, I *recognise* myself in this object-self revealed by the look of the other. Only this explains how, in the keyhole case, I come to experience shame. Thus the existence of my shame testifies immediately and unequivocally to the existence of the other.[37]

In this confrontation with the other two 'situations' collide, and each of these conflicting situations implies a different 'world'. The other cannot turn up in my world as she is in herself, nor can I turn up in hers. But we nevertheless feature as objects in each other's worlds, and we may be acutely aware of the various ways in which we do so. With the advent of the other I experience the 'draining away' of my own world.[38] This draining away occurs as the effect of the other's 'look'. If I have the strength, I may be able to turn the tables, and turn my own look upon the other. Thus the possibility arises of a battle of looks, and of worlds, in which, while neither party can ever be fully present in the other's world, we are each aware that we feature as a recognisable *object-self* (not necessarily a pure body!) in the other's world.[39]

With this account of the experience of the other, Sartre has not set out to disprove solipsism. Nor has he set out to prove that being-with is a *necessary* element in human consciousness. Rather, the point is that, mysterious though it may be (given the empirical inaccessibility of the other as a self-present duality), I cannot as a matter of fact doubt that the other impinges on my world. Consequently, I cannot in good faith deny that the other exists as a subject – a being capable of being the origin of a novel (to me) situation, and capable of taking me as her object (where this object is an object of possible recognition for me). In falling under the other's look I find myself objectified – transformed into an object self (not a mere object,

---

[35] Ibid., pp. 270–1.

[36] Ibid., p. 261.

[37] Sartre tells us: 'Shame is by nature *recognition*. I recognise that I am as the Other sees me.' Ibid., p. 222.

[38] Ibid., pp. 254–6.

[39] Ibid., p. 363.

not a non-person, lacking self-conscious goals and all 'mental' attributes – but not a self-as-self-present-duality either).[40]

I may, it seems, experience this objectification with shame, with pride, or with a range of other emotions. Nevertheless, it will, on Sartre's view, in all cases involve an element of instrumentalism – as I come to figure in the other's world in a largely instrumental fashion. Sartre's account of the instrumental objectification we suffer under the look of the other is closely linked with Heidegger's concept of inauthenticity. But there are nevertheless key differences between the two accounts. Heidegger's primary concern seems to be with the fact that, by interpreting myself along instrumental lines, I give a spurious 'fixity' to my world, denying the practical origins of the relevant instrumental complexes. Sartre's account includes these elements, but adds something further: a sense of a genuinely painful encounter with another being for whom I figure as a recognisable (to me) object. Sartre's account is thus doubly reflexive. Not only do I see myself potentially embedded in an instrumental complex, I also see my potential instrumentalisation from the perspective of the other. (Simultaneously I am aware that the tables could be turned, such that the being I recognise as another like me might be reduced to an object within the situation as *I* see it.)

Sartre repeatedly emphasises that the relation with the other is one of *being*, not one of knowledge. By this he means that it must be understood hermeneutically, as something consciousness lives through, rather than as something that is grasped purely intellectually or theoretically. Thus one does not simply know that one is subject to the look of the other. The other's look is something of which we have a more immediate, intimate and troubling awareness. That said though, the relationship with the other certainly *involves* knowledge, for it centrally involves an awareness of myself as an object of knowledge for the other. My awareness must involve a sense of what the other is likely to be *making* of me, else I would have no reason to react in any particular way – with shame, pride etc. If, in falling under the other's look, I were simply to become aware of the other's 'awareness' of me, without this more developed cognitive element, my experience could not evolve the degree of specificity required to provoke in me a reaction of shame (or pride).

9.4 Modes of Instrumentalisation

The previous section outlined a Sartrean conception of objectification as a matter of falling under the other's look, and thereby coming to figure as a (recognisable) object-self in an alien world. But while it has been shown that my objectification at the other's hands may potentially be painful for me (as in the case of shame), it has not yet been shown that my objectification is *necessarily* something negative. Indeed, since pride involves objectification just as shame does, it may presumably

---

[40]Ibid., pp. 262–3.

even be pleasurable.[41] What then is *wrong* with being objectified, if the experience may turn out to be pleasurable? To answer this question we need to turn to Sartre's account of concrete relations with others.

9.4.1 Indifference

Thus far we have a picture of relations with others as a matter of a battle of looks, and attempts at mutual objectification. This picture is consistent with much of what Sartre has to say about concrete relations with others. He is happy to use the language of conflict and mutual attempted 'enslavement' of the other. Indeed, he is prepared to state that, 'conflict is the original meaning of being-for-others'.[42] What this simple statement fails to do however is to bring out the potential complexity of the relation with another. Beyond the simple battle of mutual attempted instrumentalisation, there proves to be a more complex level, on which a different type of conflict takes place, a conflict having the character of a 'game of mirrors'.[43]

Sartre terms one common mode of concrete relation with others 'indifference'.[44] Indifference is associated with a radical instrumentalisation of the other, in which she is reduced to a mere functional unit, and never really features as a *person* for us at all.[45] Although indifference is a matter of the assertion of one's own subjectivity, and the associated collapse of the subjectivity of the other, indifference never really registers the other's subjectivity. This mode of indifference to the other is responsible, we are told, for the rise of an entire literary genre exploring the possibility of the Machiavellian manipulation of the other.[46] It is marked by an apparently impregnable confidence in the face of others, and by a noteworthy blindness with respect to them. The state of indifference may, it seems, be momentary; or it may be life-long. Some people, Sartre observes, 'die without – save for brief and terrifying flashes of illumination – ever having suspected what the Other is'.[47]

The pure instrumentalisation of the other associated with indifference conforms closely to some aspects of Heidegger's concept of inauthenticity. However, Sartre is clear that indifference by no means represents the typical mode of human interaction with others. My initial discussion of shame and the look may have suggested a relatively simple picture of a battle of straightforward mutual instrumentalisation,

---

[41] Ibid., pp. 263, 377.

[42] Ibid., p. 364.

[43] Ibid., p. 376. Sartre ascribes this character specifically to love relationships, but it is at the same time clear that his description of love relationships is intended to shed light on inter-personal relationships in general.

[44] For ease of exposition I am departing from Sartre's own order of presentation here. It is worth noting that he acknowledges (Sartre 1989, p. 364) that there is no necessary order of presentation for these modes.

[45] Sartre 1989, pp. 380–1.

[46] Ibid., p. 381.

[47] Ibid.

but that is not how human beings really behave, for the most part. As a human being, Sartre suggests, it is practically impossible to be systematically indifferent to others. Indifference is in reality always troubled, haunted by a sense of 'lack and uneasiness'.[48] This uneasiness has its roots in the evident connection between the awareness of the other, and objectivity.

We noted in the previous section that the importance of our sense of being-with-others lies partly in the way that some such sense founds our notion of objectivity – for example, scientific objectivity. The very idea of objectivity implies availability-to-others (in principle). Thus if we had no fundamental awareness of the other, we would correspondingly lack all sense of objectivity. Openness to the existence of the other is therefore bound up with the possibility of objectivity with respect to our own self-understanding. We cannot live our lives in good faith without some sense of objectivity, and for this we are ultimately dependent on the other.[49] To be deprived of any concept of the other would be to be deprived of the possibility of meaningful self-knowledge. Concomitantly, the unease of the indifferent individual originates in, and betrays, a failure to acknowledge her own fundamental dependence on the other. The power that the indifferent individual seems to have over others is therefore double-edged. It is a power that simultaneously makes it impossible for the indifferent individual to live a 'self-possessed' life, aware of her true qualities, her talents, and their limits. Being incompatible with genuine self-knowledge, it is incompatible with any genuine sense of self-worth.

9.4.2 Love, and Love's Fragility

In Sartre's view, most of us, most of the time, are not indifferent to others. The brutal instrumentalising treatment of others associated with indifference is a recognisable form of human behaviour, but it is certainly not typical human behaviour. More typical of our modes of concrete relations with others is behaviour that at some level implicitly acknowledges our dependence on the other, for our own self-knowledge, and sense of self-worth. These more typical modes do not however lift us into a sphere that is free of conflict. Rather, they lift us into a mode of relationship to others in which the conflicts that typify our interactions take a more complex form.[50]

If Sartre's enquiry into concrete relations with others stopped short with the account of indifference, he would have painted a picture of human relationships as essentially involving conflict, and as essentially instrumental. This picture would however be devoid of some of the more fascinating elements of such relations – of the kinds that occur in, for example, sexual relationships and love relationships.

---

[48]Ibid.

[49]It is, Sartre tells us, Hegel's brilliant intuition that I 'depend on the Other in my being' (Sartre 1989, p. 237).

[50]Specifically, that of the 'game of mirrors', referred to earlier.

Thus it would be radically inconsistent with what literature and the arts, high culture and popular culture, tell us of human life – where sexual and love relationships seem to be central.[51] Sartre certainly considers love and sexual desire to be central to our concrete relations with others.[52] He does not however view intimate relationships with others as providing any kind of refuge from the basic situation of conflict. Indeed, it seems that for him the basic forms of conflict receive their clearest expression in sexual and love relationships.[53]

If love is central to human existence, what is it we want when we wish to be loved? Sartre claims, plausibly I think, that one of the things we seek is the objectivity of the other. We want the external perspective that helps bring us to self-knowledge, and a reliable sense of self-worth.[54] There is therefore an original impetus of good faith in the search for love. We want something from the other that we cannot supply for ourselves: an independent perspective on ourselves. We are not content to remain within the self-referential world of the indifferent, with its characteristic unease.

It follows that what we seek in a love relationship is something analogous to our own objectification. Needless to say, this is a risky project. The external perspective will not necessarily be kind. And even if kind, it will not necessarily bring good news. Why then do we seek it? Sartre suggests that the one who seeks love doesn't typically demand that the other paint an attractive picture of her. If the motive that distinguishes love from indifference is the aim, consistent with good faith, of obtaining an objective view of the self, the demand that the lover whisper only flattering endearments is incompatible with the search for love. Much more important, according to Sartre, is the demand that the lover treat us as an absolute, and as a limit to her own projects.[55] That is to say, whatever interpretation of us might issue from the lover's perspective, the typical demand that love makes is to be loved in spite of everything. However painful the portrait the lover paints, we still demand that she recognise our ends as her ends, that we should be her all-in-all, and a potential limit on her personal projects. We don't flee from her truth, and in that sense we don't demand idealisation. But we do demand unconditional commitment – to be

---

[51] On the other hand, the portrayal of indifference as typical would be reminiscent of certain political-psychological accounts of human nature that have originated from apologists for both liberal capitalism and illiberal forms of totalitarianism. The portrayal of concrete human relations as essentially taking the form of indifference can never easily be made plausible, nor can it be considered ideologically neutral.

[52] In criticism of Heidegger, who doesn't make the slightest allusion to sexuality in his existential analytic, Sartre asks, 'Can we admit that this tremendous matter of the sexual life comes as a kind of addition to the human condition?' (Sartre 1989, p. 383).

[53] In focusing on sexual desire in particular as key to the analysis of concrete relations with others, Sartre may seem to be revealing a basic Freudianism. However, Sartre rejects Freudianism (Sartre 1989, p. 54), and must do so, given his commitment to the doctrine of the transparency of consciousness, which allows no role for the unconscious.

[54] Sartre 1989, pp. 365–71.

[55] Ibid., pp. 368–9.

treated as an unsurpassable end.[56] To this end the quest for love becomes the quest to fascinate another. We seek to become an object for the other, but an object of a very special kind: an unsurpassable object, an object that can never be a means, but only ever an end. This requires that the other be fascinated, since we cannot produce the desired result except by engaging her freedom – the love relation must be freely entered into.[57]

Consequent upon these various factors, the state of love is inherently fragile and unstable. It is constantly prone to collapse, and this collapse typically occurs in one of three ways:

Firstly, love may collapse through the instrumentalisation of the lover. The quest to have another 'objectify' oneself requires a delicate touch with respect to the other's freedom. Compulsion or coercion of the lover is a constant temptation. But it brings in its train the collapse of love, as one partner effectively dominates and instrumentalises the other.[58]

Secondly, perhaps from fear of bringing about the collapse of the relationship through coercion of the lover, there is the possibility of giving in to the opposite temptation: that of surrendering oneself entirely to the other, as an object to be used or abused – one gives up one's claim to the status of absolute end, and settles for the instrumental objectification of oneself. This also spells the end of the love relation, though this time it ends in masochism.[59]

Finally, even if one is able to maintain, in the context of an intimate relationship with a lover, the required delicate balance between avoiding encroaching on her freedom, and surrendering entirely to the other as a means, there remains the constant danger that both partners will suffer objectification by some third party. The precious transcendence of the other, which is the foundation of the desired objective perspective on myself, and which I seek to captivate without coercing, may be annihilated in the presence of a third party. Under the look of a third party my desired absolute foundation may suddenly become relativised. The hitherto 'transcendent' lover is now herself a 'transcendence transcended'; and the sought-for love relation is transformed into a relation between two mere object-selves.[60]

Thus love, as a fundamental mode of being-for-others, contains the seeds of its own destruction.[61] It may collapse into a relation of domination of the beloved by the lover; or it may collapse into masochism, in which the beloved courts domination by the lover. Finally, it may collapse through the joint objectification of the couple by a third party. This 'triple destructibility' of love does not however exhaust

---

[56]Ibid., p. 369. Sartre remarks that if only we could interiorize the entire system the lover-beloved would have become its own foundation (Sartre 1989, p. 371).

[57]Ibid., pp. 364, 371–2. Of course, although we wish to know ourselves through another's eyes, we do not (initially at least) seek our own instrumentalisation at the hands of the other.

[58]Ibid., p. 376.

[59]Ibid., pp. 377–9.

[60]Ibid., pp. 376–7.

[61]Ibid.

the routes by which a fundamental mode of our being-for-others may collapse into domination or masochism. Having described love's destructibility, Sartre observes that, 'One would have to be alone in the world with the beloved for love to preserve its character as a fundamental axis of reference', before adding, rather curiously, 'hence the lover's perpetual shame (or pride – which here amounts to the same thing)'.[62] Here Sartre seems to suggest that the nature of love, as a mode of paradoxical self-objectification, eagerly sought yet permanently exposed to collapse and exploitation, somehow combines, or makes indistinguishable, the emotions previously connected with the cruder version of the conflict of looks, discussed earlier. He seems to be at least hinting that not only does the emotion of shame have the power to reveal the modes of our being for others (somewhat as, for Heidegger, anxiety has a world-revealing power), but it is at the same time intrinsically connected with modes of objectification, as if these were occasions for shame, or a mysterious shame/pride combination. We will follow up this hint shortly. For the present I want to turn to consider another fundamental mode of our being for others that also typically proves to be a gateway into domination and masochism: sexual desire.

9.4.3 Sadism and Sexual Desire

Sexual desire, Sartre notes, is not reducible to a feeling. Sexual desire may be associated with a feeling, or feelings, but no mere feeling could have the consequences of sexual desire.[63] When we are suffused with sexual desire we desire something specific. A mood, such as melancholy, typically is not directed toward a given 'object', it is more a basic (if temporary) orientation toward the world in general. But while sexual desire may temporarily 'colour' my existence, it is nonetheless directed toward something specific; and the character of this directedness-toward explains the nature of the associated feeling(s), rather than vice versa. Moreover, unlike hunger, for example, sexual desire is not desire for a mere physical object. Sexual desire is directed toward an object, but the 'object' in question is not just an object. It is not desire for something merely physical, since it is desire for another person; but even then, it is not merely 'desire-for-another-person'. Indifference represents an illuminating contrast case here. It is perhaps possible to imagine an indifferent person experiencing desire for intercourse with another human purely as a physical object. Most of us would however be reluctant to think of this as genuine sexual desire. It is at the very least sexual desire perverted, or malformed. One may, in a sense, satisfy sexual desire with the use of a mere object. But it does not follow from the fact that sexual desire may be 'satisfied' in a variety of ways that any object that is a means to satisfying sexual desire is the *proper* object of sexual desire (this again argues for the insight that sexual desire is no mere feeling). Quite independent of

---

[62] Ibid., p. 377.

[63] Ibid., pp. 384–5.

any moralising view, it seems evident that a mere object can be at best a substitute for the proper object of sexual desire. There is then, Sartre suggests, something in sexual desire proper that amounts to desire for the other as more than just an object. Sexual desire is desire for the other *as* other, and thus as more than just a physical (or physical-plus-mental) object. Desire aims at possessing the other's transcendence – at possessing the other as self-present duality.[64] It is therefore an extremely significant phenomenon, from Sartre's standpoint (and the fact that it has been largely overlooked by previous existentialist philosophers begins to look very revealing). Sexual desire itself testifies (as did shame) to a residual awareness of the other, as something other than a mere object. But how, if to bring the other within my world is to reduce her to the status of an object, can sexual desire ever be satisfied? How is it possible to 'possess', even momentarily, the other's transcendence?

The key to the possibility of possessing the other's transcendence, on Sartre's account, is the suspension of my typical instrumental modes of relation to others, and objects within my world.[65] Various modes of 'relation' of one's consciousness to one's physical self are possible. One can 'animate' one's physical self actively, as 'body', when one engages in some physically demanding goal-oriented activity.[66] But one can also animate one's physical self more passively, as in aesthetic appreciation, where transcendence is equally present, but is incarnated (Sartre says), as 'flesh', rather than body. And it is as incarnate in the mode of 'flesh' that the other's transcendence is available for 'possession'.[67] For it to be possible for the other's transcendence to be possessed however, it is necessary that both parties become flesh. If I were to be present as 'body' to the other's 'flesh' then (since mine is an instrumental mode of engagement) her transcendence would withdraw, just as when she is the focus of my 'look'. The other's transcendence can only be 'possessed' or captured by the more passive route that leads through my also becoming flesh. Desire therefore becomes the project of becoming flesh for/to the other's flesh.[68] Sartre is careful to describe sexual desire, thus understood, in rather neutral terms. The condition of being flesh for/to the other's flesh may (possibly!) sound attractive, but to experience it as overtly desirable would presumably be to come too close to the instrumentality whose (partial) suspension is the hallmark of becoming-flesh.

Desire, thus construed, is never held up by Sartre as an ideal. Nor does he suggest that its goal is attained with any great frequency. Instead, it seems, the primary significance of desire lies in the consequences of its failure. The failure of desire is, Sartre suggests, far more common than its success. It can fail in one of two ways, both connected with the *pleasures* of fleshly co-presence. On the one hand I may respond so greedily to the pleasure I obtain from the other that I become predatory

---

[64] Ibid., p. 394.

[65] Ibid., pp. 385–9.

[66] Ibid., pp. 320–1.

[67] Ibid., pp. 388–91.

[68] Ibid., p. 391.

and sadistic – utilising her as an instrument for my own pleasure. (Sartre characterises this response in terms that fit the traditional male sexual role, in which the aggressive sexual appropriation and penetration of the other result in my greeting her flesh not with my own flesh but with body – an instrumental mode of incarnation.) The other's body is then:

> [S]urpassed toward its potentialities [and] falls from the level of flesh to the level of pure object.[69]

Alternatively, my own pleasure may cause me to turn in upon myself, masochistically. I break off the mutuality of the relation, and try to elicit an instrumentalising response from the other. I remain flesh, but bring about the other's transformation into body. (This sort of response is characterised by Sartre in terms that fit the traditional female sexual role, in which the partner provoking the instrumentalising response is essentially passive, compliant and accessible.) What results is a '[demand] to be apprehended and transcended as body-for-the-Other'.[70] I become an object for the other's look. Thus the collapse of love, discussed in the previous section, is not the only route into masochism. There is another route into masochism from the failure of sexual desire to achieve the fleshly co-presence at which it aims. Nor is the collapse of love the only route into dominating/instrumentalising relationships with others. The collapse of sexual desire may also lead, as we have seen, into sadism.

These results reinforce the general picture Sartre paints of concrete relations with others, according to which they are marked by a restless circulation between domination and submission, sadism and masochism – with love and sexual desire, each promising their distinctive form of fulfilment, constantly drawing us into unstable modes of relation with others, which then repeatedly collapse into something much more negative. In all of these cases, the failure of the projects of love and desire is marked by an upsurge of instrumentalism. That each collapse represents a *failure* though implies that love and desire remain, often unrecognised, the fundamental motivating goals in sadistic and masochistic behaviour.

As I have characterised the above behaviours, they may appear to be essentially tied to love and desire, as if they could only arise through the decay of these particular forms. But in fact, Sartre suggests, although masochism and sadism ultimately aim at love and desire, their coming into being is in no way dependent on even the partial fulfilment of these goals. Typically, it would seem, love and desire fail almost from the start, lapsing into sadism or masochism before anything like love, or genuine fleshly co-presence, is even on the horizon.[71] Thus the typical modes of relation to others are various modes of failure, and the bleakness of Sartre's view is reinforced. Intimate human relationships are, it seems, predominantly sadomasochistic – albeit for reasons that may not always be fully appreciated.

---

[69] Ibid., p. 398.

[70] Ibid., p. 397.

[71] Ibid., pp. 405–8.

9.5 Non-complementarity

Sartre's account of concrete relations with others may appear to be marked, from what has been said to this point, by a fundamental complementarity. If there is indeed a powerful tendency for human relationships to fall into one or other of a set of recognisably sadomasochistic patterns, then it might well be thought that these patterns themselves will tend to display a certain stability and balance, such that one partner's characteristic sadism will be matched by the other's masochism. Any such sadomasochistic coupling will be asymmetrical, in the sense that each partner plays a fundamentally different role. But the roles will nevertheless be 'matched' – they will dovetail together – and they might therefore be characterised as complementary. Nevertheless, it should be noted that nothing in Sartre's analysis requires this sort of complementarity between sadistic and masochistic roles. Sartre does not dwell on this point, but its independent importance seems to require that we pay some attention to it.

We have seen that the project of possessing another's transcendence may fail by falling into a sadomasochistic coupling. But nothing in Sartre's analysis seems to rule out a situation in which the project fails purely as a consequence of an upsurge of sadism on the part of both parties. In such a case, the parties would lapse into mutual instrumentalisation, and might do so more or less simultaneously. (It is for example easy to imagine how this might happen following the collapse of a love relationship, taking the third of the forms that Sartre distinguishes – in which both parties experience a mutual objectification at the hands of a third party. Once the third party has left the scene (perhaps our couple are now travelling home together, and a heavy silence reigns…), it may prove impossible for either party to recover their previous non-instrumental attitude to the other.)

That said, a failure of complementarity need not take the form of a collapse into mutual instrumentalisation. We have already seen that the conflict that Sartre takes to be characteristic of human relations can take the more simple form of mutual indifference. In a conflict marked by mutual indifference, the element of attempted instrumentalisation would still be present, though in this case there would be no attempt to 'possess' the other's transcendence (the existence of which is not even suspected). Moreover, while both indifference and sadism may be met with masochism, they surely don't require it. One may be the victim of another's indifference or sadism without in any way embracing it, or revelling in it. Sartre characterises masochism as a vice, and vice in general as a love of failure.[72] But nothing in his analysis entails that when sexual desire or love collapse into some form of domination, or some struggle for domination, the victim must 'love' or in any way find pleasure in that failure.

We should beware then of attempting to impose any overly-neat or symmetrical pattern onto the failures of love and desire Sartre outlines. Where the collapse of desire involves a struggle for domination, this may, it seems, take either a one-sided

---

[72]Ibid., p. 379.

or a two-sided form, and may be marked either by the sadism of one or both parties, or by the mere indifference of one or both. Again, insofar as a one-sided domination is associated with, or elicits, a masochistic response, this response may arise in the face of either sadism or indifference. While in each case the collapse is marked by some form of upsurge of instrumentalism and domination, that upsurge may or may not involve a struggle for domination, may or may not involve sadism, and may or may not be associated with a masochistic response.

We already have a clue as to how indifference is to be distinguished from sadism. As noted previously, indifference does not partake of the 'game of mirrors'. It takes the form of an instrumentalisation of the other, but it does not seek the other's objectification – in the full sense of reducing the other to an object *in her own eyes*. Consider now, by contrast, a situation in which love collapses according to the second mode that Sartre identifies – in which the beloved, from anxiety not to destroy love by instrumentalising the other, permits herself to be instrumentalised, and falls into masochism. Here the masochistic party *still wishes to play the game of mirrors*, because she still desires her objectification at the hands of her lover. But unless we assume that the other party simply cannot fall into the type of bad faith associated with indifference – and, aside from implying that genuine indifference is highly uncommon, Sartre gives us no reason to assume this – there seems to be nothing to stop him from opting out of the game of mirrors altogether at this point, and falling into an 'indifferent' instrumentalisation of his partner. Thus a masochism/indifference pairing looks perfectly possible, and in this case the objectification of the masochistic party, while actively sought by the masochist, would be a matter of indifference to her partner, who aims only at her instrumentalisation – he is, precisely, indifferent to the matter of what she 'means' to him, or to herself – he simply wishes to use her.

Having clarified this possibility, it is apparent how sadism may be distinguished from indifference. The sadist, like the masochist, *wishes to remain within the game of mirrors*. That is, he aims not only to dominate the beloved by instrumentalising her, but also to maintain her objectification. He wishes to use his partner, but also to maintain her perception of herself as an object, being used. One cannot play the sadist's part, then, without remaining within the game of mirrors; though one can certainly maintain indifference outside of the game (and this is in turn the origin of the uneasiness and bad faith that haunts indifference). While indifference only seeks the other's instrumentalisation, sadism utilises the other's desire for objectification to give an additional twist to the knife, an additional element of domination. Indifference wants a tool. Sadism wants a tool that knows itself as tool – a subject-tool. Accordingly, a key feature of Sartrean sadism, which is missing in the case of indifference, is the element of *mendacity* that sadism involves. Sartrean sadism involves *knowingly* undermining the personhood of the other – where what is known is not simply that the other is being used as an object, but also that the objectified and instrumentalised other is in reality not a mere object at all. And, of course, this element of mendacity, which serves to distinguish sadism from mere indifference, is also a feature of masochism. Sadism may or may not be associated with a masochistic response. But where it does receive such a response, the victim will be complicit

in her own objectification. Masochism involves knowingly undermining one's own personhood, in a manner that is as mendacious in its own way as the conduct that is characteristic of sadism.

Thus while Sartre's general approach in his analysis of concrete relations with others shares many similarities with Heidegger's existential analytic, his findings represent a clear advance over those of Heidegger, for our purposes. Sartre pays a good deal of attention to the doubly reflexive nature of human relations with others. These relations are not simply a matter of what we do, and of what we make of what we are, on the basis of what we do. They unfold within the game of mirrors, in such a way that one or other party may come to 'taste' their instrumentalisation at the hands of the other. It follows that, where instrumental stereotyping is in evidence, attention should be paid not simply to instrumentalising conduct, and inauthentically instrumentalising interpretations of oneself and others, but also to the different ways in which individuals may respond to such instrumentalising interpretations. The undermining of another's personhood may be essentially 'indifferent', or it may be knowing, and mendacious. It is the latter form that Sartre regards as characteristic of sadism.[73]

Furthermore, while Heidegger's analysis of being-with others as a dimension of our being-in-the-world is relatively generalised, Sartre plunges into a detailed investigation of the most highly-charged and intimate areas of our interactions with others, uncovering a range of phenomena that Heidegger's analysis of inauthenticity fails to even hint at. While it is evident that the sort of conduct that Sartre regards as sadomasochistic would qualify as inauthentic in Heidegger's terms, involving as it does inappropriately instrumental interpretations of oneself or another, it is also evident that Sartre's analysis, by uncovering the element of mendacity in both sadistic and masochistic conduct, picks out forms of inauthenticity that promise to have a good deal more immediate ethical relevance. Sartrean indifference, whilst being a particularly brutal mode of relation to others, preserves a measure of relative innocence, due to its lack of the element of mendacity. But genuine indifference is, according to Sartre, a highly untypical state. Predominantly, for Sartre, human relations are played out in the realm of the modes of objectification associated with sadism and masochism.

Sartre's ultimate conclusion then is that both intimate personal relationships, and more general social relationships, are distinctly sadistic, or sadomasochistic. The general pessimism of the view is relieved slightly by the fact that the slide into sadism, or sadomasochism, is initially motivated by love and/or sexual desire; and love is motivated in part by a desire for objective self-understanding. On Sartre's account, genuine authenticity demands objectivity in respect of oneself: one must be aware of, and accept, one's facticity. We are reliant on the other for any such objectivity, in two distinct ways. Firstly, and most obviously, others provide us

---

[73]Thus, if it needed underlining, we cannot necessarily regard Sartre's analysis as an analysis of sadism and masochism in general, as conventionally understood – since these can apparently take the form of a quite straightforward enjoyment in inflicting or suffering pain, without the doubly reflexive element associated with Sartre's 'game of mirrors'.

with alternative perspectives on ourselves – views of who we are, and what we do, which originate independently of our own ego, and can therefore make some claim to objectivity. Secondly, and more fundamentally, our very *idea* of objectivity requires us to have some awareness of the existence of the other. To live in a genuinely solipsistic manner would not only be to live without access to an alternative viewpoint on ourselves, it would be to live without acknowledging such a viewpoint as a real possibility. It is the drive for objectivity with respect to ourselves – to capture the secret of our own being, which the other is seen to possess – that, at least in large part, leads us to seek love, and which, in turn, precipitates the collapse into sadomasochism.[74] Not only does the desire for objectivity with respect to oneself provide the motive force that keeps the wheel of love-desire-sadism-masochism turning, it also contributes the element of mendacity that is a characteristic feature of Sartrean sadomasochism. For Sartre, neither sadism nor masochism would be what they are unless, at some level, they acknowledged what they simultaneously deny: the personhood of the other. Moreover, the sadist does not simply knowingly deny the personhood of the other (nor does the masochist simply knowingly deny her own personhood). The sadist aims to humiliate his victim – to ensure that she 'tastes' her own objecthood. (The masochist, correspondingly, seeks her own humiliation – to taste her own objecthood.) But one cannot humiliate a mere instrument. One can only humiliate an ontological person – an individual capable of maintaining a sophisticated self-conception of her own, and, potentially, of suffering by it. Thus Sartrean sadism and sadomasochism are thoroughly mendacious modes of relation to others.

This suggests the possibility that a revised ethic of respect for persons might incorporate the sphere of interpretive moral wrongs while at the same time being couched in terms of public obligations, rather than simply in terms of private perfectionism. Both the indifferent individual, and the sadist, are guilty of the instrumental stereotyping of their respective victims. However, the culpability of the indifferent individual seems relatively mild, given the corresponding innocence that attaches to his state. It would be no easier to argue that we have a perfect duty not to be (genuinely) indifferent toward others than it would be to argue that we have a general duty to be authentic. To be either indifferent or inauthentic would simply be to fail to actualise certain important human possibilities – generally akin to a failure to develop one's talents. But, whether or not Sartre would have been interested in developing this line of thought, it is certainly possible to argue that we have a moral duty not to behave (in his terms) sadistically, or masochistically. Above and beyond any harms that may be associated with such behaviour, their victims suffer a distinct wrong, as a consequence of the systematic mendacity on which such behaviour is based.

---

[74]By identifying a motive for the picture of cyclical sadomasochism he describes, Sartre's account seems to represent a further advance on that of Heidegger, for whom (the mysterious call of 'conscience' aside) the quest for authenticity seems to imply a motiveless 'leap'.

Sartre's analysis of concrete relations with others thus potentially lifts a sceptical essentialist ethic of respect for persons out of the arena of private perfectionism, and into a well-recognised area of the morality of public obligations. To engage in the instrumental and/or ideological stereotyping of persons (oneself or another) – to 'objectify' them – is, when it takes the form of a knowing restriction or undermining of personhood, as in Sartrean sadism and sadomasochism, to subject them to an interpretive moral wrong. Not only is such conduct fundamentally dishonest, but the particular form of dishonesty involved is calculated to humiliate. Both to practice mendacity with respect to others, and to seek to humiliate them, are forms of conduct widely recognised to be incompatible with respecting their dignity. The distinctive value of Sartre's analysis lies in the fact that it highlights forms of mendacity and of humiliation that an approach less sensitive to the prevalence and the significance of (what I have termed) interpretive moral wrongs might easily overlook.

References

Sartre, Jean-Paul. 1948. *Existentialism and humanism* (trans: Mairet, P.). London: Methuen.
Sartre, Jean-Paul. 1989. *Being and nothingness* (trans: Barnes, H.). London: Routledge.

Chapter 10
Foucault and Subjection

Abstract A sceptical essentialist ethic of self-interpretation, founded on an obligation to avoid the mendacity involved in inducing deficient self-conceptions in others, looks to have significant normative force. But how might it apply outside of the personal relationships investigated by Sartre, in a broader social context, in which self-conscious sadism (and masochism) seems to be uncommon? This chapter addresses this question with reference to the work of Michel Foucault. Although Foucault rejected key elements of phenomenology, his account of the power effects of disciplinary technologies has clear parallels with Sartre's account of sadism in concrete relations with others. At the same time, he emphasises that disciplinary power does not require an agent, and may be diffused throughout social institutions. Foucault did not regard himself as an ethicist, in any conventional sense; but in highlighting the price we pay for scientific self-knowledge, his findings have clear implications for those whose professional roles involve the acquisition and deployment of such knowledge.

10.1 Self-Knowledge and Power

The previous chapter closed with discussion of how an ethic of respect for persons might be developed along sceptical essentialist lines as an ethic of self-interpretation. There remain problems with such an ethic however, as developed to this point. To bring out these problems it is helpful to consider its possible applications.

In *The Second Sex* Simone de Beauvoir shows how a broadly Sartrean analysis can be transferred from the context of the personal relationships on which Sartre focuses to the level of large-scale gender politics.[1] In particular she draws attention to the pervasive treatment of women as 'other' – the negative counterpart of the male 'subject' class.[2] Even women tend to refer to women from an external perspective, speaking and writing about women as a 'they' rather than a 'we'. The apparent unavailability of the perspective of a female subject looks to be the product of a generalised male objectification of women, whose more or less 'masochistic'

---

[1] See de Beauvoir 1952.

[2] Ibid., Introduction and passim.

response to such treatment de Beauvoir considers at length.[3] However, elevating the Sartrean model to this level of generality also exposes some shortcomings. We can highlight these by considering that if the recent history of gender relations has evidenced a general male objectification of women, sticking closely to the Sartrean model would require us to explain this objectification in terms of either a general male indifference to women, or (more plausibly) a general male sadism. The problem with this is that, for all its helpfulness in other respects, the Sartrean model, raised to this level, looks both excessively pessimistic, and excessively voluntaristic. Let me briefly explain what I mean by this.

The excessive pessimism of the Sartrean model comes through when we consider what, practically, might be done about the male objectification of women, if male sadism really is as ubiquitous as de Beauvoir and Sartre suggest. Sartre treats genuine indifference as if it were, except perhaps in the most unusual cases, only a theoretical possibility. In the more usual case what looks like indifference will, it seems, turn out to be more or less well-disguised sadism. No doubt Sartre had his reasons for portraying human relations, and particularly gender relations, as fundamentally sadomasochistic. But on his presentation sadomasochism looks practically universal. It may be that for some purposes such a sweeping view is appropriate, despite – or even, perhaps, because of – the grimness of the picture it paints. But at the same time, if we are to make a syndrome the object of an obligation, it has to be evident that it is in our power to do something about it. (In the Kantian phrase, *ought* implies *can*.) Unfortunately, the Sartrean analysis seems in danger of pulling the rug from under an ethic of self-interpretation, by inviting the objection that if objectification is really such a ubiquitous feature of human relationships, it will be beyond the power of any ethical or political movement to combat. Moreover, to imply that gender relations under current conditions are characterised by a practically universal male sadism towards women would be to risk depriving ourselves of the means by which to pick out instances of *notable* male sadism.

The second problem is related to this first. On the one hand the Sartrean model suggests that practically universal objectification is more or less unavoidable, implying that an *ethic* of non-objectification would be effectively redundant. On the other hand, it seems to suggest something that hardly seems compatible with this idea: that combating objectification is a highly voluntaristic affair. If we ask how it might be possible to break the sadomasochistic cycle, the Sartrean answer seems to be: through an exercise of freedom – one must simply refuse to objectify the other.

These (linked, but in some respects contradictory) problems presumably explain, to some degree, the apparent practical ineffectiveness of a Sartrean concept of objectification at the ethical-political level. It is possible to use Sartre's analysis to identify such large-scale syndromes as a general male objectification of women. But such analyses tend to defeat themselves by their own ambitions. Supposing there is a practically universal male objectification of women? What, bearing in mind its universality, might be done about it?

---

[3]Ibid., pp. 650–1, and passim.

De Beauvoir's analysis concentrates, in terms of its practical suggestions, on counteracting female masochism. Women must refuse their object status, and aim to become a 'subject' class. But as we have seen, there is an asymmetry about the sadomasochistic relation that suggests that while sadism may be met with masochism (objectification may be welcomed – which, of course, makes it no less wrong or harmful), sadism doesn't require a masochistic response. Supposing that de Beauvoir's analysis were accurate then, the universal rejection by women of the masochistic role she implicitly ascribes to them (which some might argue has begun to occur during the period since she wrote) would not necessarily lessen male sadism in any way. There seems no reason to think, on a Sartrean account, that the sadist would be any less sadistic if he were no longer to encounter a masochistic response.

We therefore need to consider whether an ethic of self-interpretation might be developed in a manner that is able to construe objectification as a more localised phenomenon (though without ruling out the possibility of the large-scale objectification of entire social groups), and in a manner that is able to construe the struggle against objectification in a less voluntaristic manner – such that it does not simply involve individual acts of will, which apparently engage with no other mechanism (while at the same time, of course, not ruling out the perpetrator's responsibility for objectifying others). For both of these aspects I want to turn to my last major source for a contemporary ethic of self-interpretation: Michel Foucault.

It may seem strange to move from a discussion of the work of existential phenomenologists like Heidegger and Sartre to a discussion of Foucault. Foucault belongs to a group of mid- to late- twentieth-century thinkers whose philosophical point of departure is the critique of the residual Cartesianism that is a notable feature of Husserlian transcendental phenomenology, and which remains a strong influence on Sartre (less so on Heidegger). Partly under the influence of Nietzsche, Foucault and his contemporaries reacted in particular against the idea of the meaning-constituting phenomenological subject, who is blessed with perfect transparency with respect to her own mental contents.[4] Nevertheless, Foucault was not simply reacting against the phenomenological orientation of an earlier generation. For one thing, Sartre's own philosophy embodies some (recognisably Nietzschean) non- or anti-phenomenological aspects. Sartre's own intellectual trajectory led from phenomenology via existentialism to Marxism, and it would not be wholly inaccurate to say that existential Marxism functioned for him as something of an escape route from phenomenology.[5]

In addition, there are a number of common themes in the work of Sartre and Foucault. Indeed it would be surprising if this were not so, given the debts that both of them acknowledge to the work of Heidegger.[6] It is important then not to

---

[4]See Foucault 1988a.

[5]Ibid., p. 21.

[6]In an interview first published in 1984 (English translation 1985) Foucault goes so far as to say that 'My entire philosophical development was determined by my reading of Heidegger' (Foucault 1988b, p. 250). On this general theme see Lucas 2002.

exaggerate the extent to which the work of Foucault and his contemporaries represents a break with what had gone before. This chapter will not attempt to examine Foucault's relation to Sartre and Heidegger in detail. Instead, the aim will be to bring out those elements in Foucault's work that develop recognisably Sartrean themes, while at the same time showing how Foucault developed those themes in his own distinctive way, in a manner that will be of service for a contemporary ethic of self-interpretation.

A survey of Foucault's writings initially gives a scattered impression (this is particularly evident in the writings from the early seventies, and is hinted at by Foucault himself in a number of places).[7] It is difficult at first sight to see his work as belonging to a continuous philosophical project. From a background in the history and philosophy of science he wrote detailed historico-philosophical studies on a range of apparently loosely-connected topics, including psychiatry, sexual psychology, criminology, biology, economics, linguistics and clinical medicine.[8] In the process he became a hugely influential figure in social and cultural theory. His work was initially associated, by others, with structuralism. When he repudiated this label he was loosely assigned to the poststructuralist camp. However, his own views, set out clearly and candidly in a series of essays and interviews given shortly before his death, suggest that he was, ultimately, able to view his own work as belonging to a coherent and unified project, which had little to do with either structuralism or poststructuralism as conventionally understood. Summarising his own philosophical project, Foucault remarks that he has been above all concerned to pose the question, 'at what price can subjects speak the truth about themselves?'[9] More specifically, he was concerned to highlight the price we have paid for developing, particularly since the closing decades of the eighteenth century, an ever-expanding corpus of scientific knowledge of humanity – thus the price humanity has paid for its development of scientific self-knowledge.

This theme is recognisably Nietzschean. However, a modern Nietzschean could hardly replicate Nietzsche's own idiosyncratic philosophical approach, and Foucault does not attempt to do so. His Nietzscheanism is instead manifest in his approach to history as *genealogy*. Foucault's project, as set out in the programmatic essay 'Nietzsche, Genealogy, History', is to document what we might refer to as a 'hidden' history of modern thought – in particular, the hidden history of modern scientific knowledge of humanity.[10] However, his aim in doing so is not, as in so many tales of historical conspiracies, to unveil via such a 'hidden history', some deeper pattern, structure, or programme underlying events as they have unfolded. Rather, it is the very opposite: to show that anything remotely resembling structure, pattern, destiny, or a necessary historical progression is lacking.[11] What Foucault aimed to reveal are

---

[7]Foucault 1980a, pp. 78–9.

[8]Foucault 1988a, p. 30.

[9]Ibid.

[10]See Foucault 1977b.

[11]Ibid., p. 142.

the thoroughly mundane origins of key aspects and ingredients of modern thought. His goal was to write the history of ideas that are not – officially – supposed to have *that* sort of history. Foucault regards this project as inherently subversive, just as Nietzsche's account of the genealogy of morality was inherently subversive.[12] In a context in which it is of considerable cultural importance that certain phenomena are considered not to have straightforwardly worldly origins, any study that shows that they do in fact have straightforwardly worldly origins can have a significantly subversive impact, without departing from the detached perspective of the historian.

Once Foucault's project is understood in this way, it becomes apparent why his work should give such a scattered impression. His enquiries must invariably lead him away from grand theorising, into historical backwaters and by-ways. Nevertheless, there is a unifying philosophical intent: awareness of how key modern ideas originated in such backwaters and by-ways will serve as an antidote to traditional notions of modern scientific self-knowledge as a product of enlightenment, reason and progress.

Such a project may initially seem light years away from Sartre's notably ahistorical analyses of supposedly core structures of human experience and consciousness. It should be borne in mind however that one distinctive feature of Sartre's 'Cartesian' approach to the analysis of fundamental aspects of human experience is his determination to stick closely to the (indubitable) appearances, rather than searching behind them for some kind of underlying pattern or structure. As we saw in relation to Sartre's analysis of concrete relations with others for example, the concern is with what, as a matter of experiential fact, we cannot possibly doubt, rather than with any attempt fundamentally to *explain* the relevant phenomena (an attempt that promises to become speculative at some point).[13]

Moreover, the theme of the paramount significance of forms of self-knowledge is common to both Sartre and Foucault. This latter fact is hardly surprising, given that, as previously mentioned, both philosophers were strongly influenced by Heidegger and Nietzsche, both of whom were centrally concerned with the self-knowledge, but had an ambivalent attitude to it. For Nietzsche, a drive to self-knowledge is a distinctively human characteristic. Honesty remains a cardinal virtue (perhaps the cardinal virtue) for Nietzsche, who to this extent at least remains a deeply moral thinker.[14] A life that does not revolve around an uncompromising thirst for self-knowledge is, in his view, barely recognisable as a human life at all.[15] At the same time however, Nietzsche recognises that the distinctively human quest for self-knowledge is a never-ending one, since humanity itself is, properly understood, a continual project of self-overcoming. At the same time as being driven to know ourselves we are (or should be) in the process of re-inventing what it means to be human. Thus,

---

[12]Ibid., pp. 144–7.

[13]Sartre 1989, p. 251.

[14]For this view of honesty, and the associated notion of a good intellectual conscience, see Nietzsche 1974, sections 110, 114, 319.

[15]Ibid., section 110.

Sisyphus-like, the search for scientific self-knowledge is both an inescapable task, and a task that cannot be completed.[16]

The paradoxical Nietzschean struggle to attain an unattainable self-knowledge has recognisable parallels in both Heidegger and Sartre. The parallels are evident in Sartre's account of the conflicts associated with our attempts to forge, protect, and overthrow conceptions of our own identity, in the face of the other. In our struggle to acquire a sense of ourselves in our facticity, we are necessarily dependent on the other, as the only possible source of objectivity. At the same time though, we cannot evade the awareness that the objectified self to which the other introduces us is not, and cannot be, the subject-object duality that we are in truth. In accounting for the inevitable failure of the quest to acquire objective self-knowledge, Sartre is clearly indebted to Heidegger, from whom he derives the conception of the human being as the being whose own being is an issue for it – whose essence is to make itself what it will be. While Nietzsche couches this struggle in epistemic terms, as a quest for self-knowledge, Heidegger pictures it in ontological terms, as the struggle to gain the sort of fundamental self-understanding that first makes self-knowledge possible. Sartre, again following Heidegger, similarly ontologises the struggle.[17]

Foucault's treatment of the struggle for self-knowledge retains elements of all three of these approaches. The Heideggerian and Sartrean ontologisation of the struggle is rejected as being insufficiently historical. It retains a hint of Hegelianism, despite Sartre's 'Cartesian' determination not to go beyond the (necessarily historical) appearances. In this respect Foucault remains much closer to Nietzsche: the focus is not on fundamental structures of consciousness, or anything equivalent to them, but on historically specific forms of scientific self-knowledge, which can be shown to possess an identifiable historical origin.[18] In this Foucault is consciously distancing himself from those aspects of Heidegger's thought that may appear to involve mythologisation and mystification. At the same time, Foucault de-intellectualises and de-individualises the historical struggle for self-understanding, in a manner that suggests that he remains close to Heidegger in some key respects. The respective attitudes of Sartre and Heidegger to individualistic considerations remain an important point of difference between them. We have already seen how the Sartrean sadomasochistic struggle plays itself out in concrete relations between specific individuals. Heidegger, by contrast, could be said to conceive of the struggle for authenticity not as a struggle that engages us as individuals, but as a struggle from which individuality potentially emerges. That is, it is not a struggle in which we seek to defend a pre-given individuality against others who would threaten it, but a struggle in which individuals strive to establish a recognisable individuality, by escaping from the banal levelling-down that characterises the inauthentic 'they-self'.[19] Heidegger's view is anti-individualistic, at least to the extent that he sees individuality not as a given but as an achievement.

---

[16]Nietzsche 1966, section 227.

[17]For example, with his notion of bad faith.

[18]See Foucault 1977b.

[19]See Heidegger 1962, p. 167.

Thus the particular combination of elements that go to make up the Foucauldian approach: an epistemic focus, combined with a refusal to present historical struggles primarily in intellectualistic terms, and an anti-individualistic perspective, add up to a distinctly pessimistic philosophy, in which the inescapable struggle to gain scientific self-knowledge is haunted by the perception that such knowledge is never neutral, and always serves some interest. Knowledge, he states bluntly, 'is not made for understanding, it is made for cutting'.[20]

In the works from the 1960s, Foucault's main concern is to identify some of the very general historical shifts that have affected the acquisition of scientific self-knowledge in fields such as medicine, psychiatry, economics, linguistics and biology. The aim here, as explicitly announced in *The Archaeology of Knowledge*, is to identify the 'rules' that, at any given time, enable a theory, an account, even a particular document, to qualify as scientific – as belonging to a scientific corpus of knowledge.[21] A particularly clear example of this approach in practice can be found in *The Birth of the Clinic*. Here Foucault is concerned with a development that affected the very foundations of medical science, but has subsequently seemed to be almost unnoticeable. This is the shift from an organic to a reductionist view of disease – from, as Foucault puts it, the perspective of a botanist, to that of a chemist.[22]

The nature of this shift can be illustrated by analogy with the following otherwise unrelated case. Consider the different ways in which a builder and a mycologist might view a rotten wooden beam in a house. To the mycologist the beam is both environment and organism. He will be conscious of being in the presence of a particular organism, in addition to the beam. He will also be conscious that the beam is an environment for that organism – supplying nutrients, moisture etc. Almost like a gardener examining a plant, he will understand the beam as a mini-ecosystem (albeit a fairly open-ended one), in which familiar biological processes are playing themselves out. The beam is the support for the organism, whose presence and life course are, in principle at least, directly observable. The builder, by contrast, will not take himself to be in the presence of anything of the sort. Rather than damp-wood-plus-organism the builder will take himself to be in the presence of a defective piece of timber. Rather than seeing a piece of timber and something more, the builder will see something less than a piece of timber – namely, timber lacking structural integrity. Given their particular practical orientations to the beam, it is evident why each of these individuals will understand the timber as they do. 'Rot' is, for a mycologist, a living presence in the world. For a builder, it is an

---

[20]Foucault 1977b, p. 154.

[21]Foucault 1989a, p. 207. Note that Foucault's 'archaeological' method, as set out in this work, is not identical to his later 'genealogical' method, referred to earlier. While the emphasis on historical origins remains a constant, the genealogical method shifts the explicit focus from the historical formation of rules, to concrete historical struggles. This requires, as Foucault expresses it in 'Truth and Power' (Foucault 1980b, p. 114), a shift from analyses couched in terms of language and signs to analyses couched in terms of force, power and tactics.

[22]Foucault 1989b, p. 119.

absence of structural integrity in timber. We do not ordinarily think of either mode of understanding as straightforwardly correct or incorrect, we simply call on them for different purposes.[23]

These two types of view have close analogues in the history of medicine. It is perfectly possible to look at a sick human body in a manner analogous to the mycologist; and indeed, according to Foucault, up to the end of the eighteenth century, the 'gaze' of the nosographer was analogous to that of a gardener. Alternatively we can approach the sick body as the builder approaches the timber. And again, this sort of view of the sick body, as essentially mechanically defective, was distinctive of the early nineteenth century. In the transition from the eighteenth century view to that of the nineteenth, the medicine of 'diseases' was replaced, according to Foucault, by the medicine of pathological reactions.[24] The sick body was no longer viewed as the environment of an organism (the disease in question), but as something less than a healthy body – something lacking what every healthy body has. The issue is not whether one view is right and the other wrong, or even whether one works and the other doesn't – for each may 'work', for certain purposes, and within certain limits. Rather, what is significant is that these alternatives exist, that they are historically distinct ways of knowing the sick human body, and that the transition from the one to the other amounted to a radical shift in the way medical science was pursued – a shift that was not independent of the broader social and historical context. Foucault shows how the shift in the nature of the medical gaze that took place between the eighteenth and nineteenth centuries was associated with the rise of pathological anatomy, and the birth of the hospital clinic, as the paradigmatic arena in which medical care is delivered. These social transformations helped make modern medical science, and in turn modern scientific self-knowledge, possible.[25]

A second illustration is given in the first volume of Foucault's *History of Sexuality*. The sexual psychology of the nineteenth and early twentieth centuries was founded on an evidence base of individual testimony. What gradually took shape in this period was, in Foucault's words, the emergence of 'a confessional science'.[26] It takes something of a re-adjustment, a critical step back, to find this idea strange. But from a historical point of view this development certainly seems to stand in need of explanation. What had hitherto served to distinguish scientific evidence from mere hearsay was the possibility of inter-subjective verifiability. And yet sexual psychology could not have arisen except on the basis of a view of first-personal accounts as a potential source of scientifically acceptable data.[27] There is a development here that cries out for explanation. It is reasonable to be curious about the origins of

---

[23] And lest it be thought that one of them is 'objective' and the other merely 'instrumental' it should be borne in mind that it may be the builder who needs to call in the mycologist for an expert opinion.

[24] Foucault 1989b, p. 191.

[25] Ibid., pp. 196–8.

[26] Foucault 1981, p. 64.

[27] Ibid., part 3.

the kinds of 'confessional' data that sexual psychology requires, and in particular about the origins of the idea that such data could qualify as scientifically valid – the more so since the fact that such a development could take place at all suggests the existence of significant links between the development of sexual psychology and the first-personal style of philosophical reflection introduced into modern thought by Descartes.[28] Foucault's analysis of the origins of the uniquely western *scientia sexualis* led him to investigate the religious institution of confession as a source of authoritative self-knowledge, and as an epistemic 'technology'.

We tend to think of the discourse produced in confessional situations as a liberating force. We unburden ourselves of deeply felt doubts, worries and convictions, both in religious confession, and to the psychoanalyst. But why think of such confessional procedures as in any way liberating or emancipatory? For Foucault, the primary significance of such confessional techniques is as mechanisms of power. It would be wrong, however, to regard confessional methods as essentially geared to intelligence-gathering. According to Foucault their connection with mechanisms of power does not take such a crude form. After all, throughout its history the religious confessional has produced a constantly vanishing archive.[29] It cannot therefore reasonably be regarded as primarily a technology for the extraction of information. Instead, Foucault argues, its primary significance lies in the way that it serves as an effective means by which to persuade the confessee to adopt a particular self-conception. In confessional procedures, one is required to produce a discourse of the self. This discourse purports to be true, and as such it has a distinctive power effect – not by supplying information to the confessor, but by being embodied into the self-understanding of the confessee. It has its power effect in the one from whom it is wrested.[30] In the individual confession, which is heavily reliant on thoroughly social notions of the acceptable and the unacceptable, thus on 'normalising' beliefs and values, it is, in effect, the Heideggerian 'they-self' that speaks. Far from being an invitation to some kind of liberation then, Foucault argues that the invitation to confess is primarily an invitation to submit ourselves to established structures of social power, by categorising and identifying ourselves in terms of what are, in effect, ideological stereotypes.[31]

Foucault's account of the origin of sexual psychology in confessional forms is thus a further example of how, rather than taking revolutions in the development and acquisition of scientific self-knowledge for granted, it is possible to account for the rise of particular scientific approaches, and particular modifications in the rules of admissibility for scientific evidence, in terms of broader cultural shifts, which at first glance may seem to have nothing to do with supposedly pure scientific knowledge. Rather than being an assortment of scattered historical surveys then, the Foucauldian corpus can be understood as a search through the history of the modern human

---

[28] See Foucault 1989b, pp. 197–9.

[29] Foucault 1981, p. 63.

[30] Ibid., p. 62.

[31] Ibid., pp. 58–63.

sciences, for evidence that will link transformations of the scientific *episteme* – the complex of rules that governs what counts as scientific admissible evidence, practice or theory in a given period – with independently identifiable social and cultural shifts.[32] The task in every case is to exhibit the thoroughly worldly and historical origin – the 'birth certificate' – of changes that might otherwise have been taken for manifestations of the stately onward progress of scientific reason.

As Foucault's work entered the nineteen seventies the emphasis on linking scientific developments to historical events was both modified and intensified. Now it was not so much a matter of conducting 'archaeological' studies of the origins of specific forms of knowledge, as of locating scientific shifts against a background of the history of power relations in particular. In *Discipline and Punish* Foucault aims to link the emergence of the science of criminology with the spectacularly rapid rise of the prison, as the paradigmatic penal institution, at the beginning of the nineteenth century.

Up to the mid-eighteenth century crime and punishment in Europe took almost medieval forms. Initially, punishment consisted of a spectacular show of force, in which the might of the sovereign was unleashed on the body of the criminal, as a salutary demonstration of overwhelming power, which would serve as a deterrent to other potential offenders.[33] Such spectacles of state violence sought to induce fear and awe, to intimidate potential offenders into compliance. But as the eighteenth century progressed they increasingly failed to function efficiently as means of social control. In fact, they proved to be a destabilising force, often provoking violent reactions from the crowds who attended executions, and inadvertently presenting criminals with public opportunities to re-cast themselves as heroes.[34]

Later in the eighteenth century such methods were gradually superseded by reforming techniques that replaced torture and violence with elaborate forms of symbolism. These new techniques sought to establish an association in the minds of potential offenders, which would suggest a natural linking of crime and punishment. By this means, it was hoped, crime and punishment would be linked in the minds of potential offenders, not through intimidation – or not purely through intimidation – but in a symbolic fashion, according to which punishment would seem both proportional and inevitable.[35] Rather than presenting punishment as simply the retribution of a superior power, these techniques would communicate a lesson: each crime has its corresponding punishment, in an independent punitive economy. The intention was to legitimate punishment. But the methods employed were complex

---

[32]The Foucauldian *episteme* is thus analogous to Thomas Kuhn's *paradigm*. Note though that whereas the Kuhnian paradigm aims to capture what distinguishes the sciences of a given period from each other, the Foucauldian *episteme* aims to capture what unites them, and at the same time distinguishes them from those of other periods. *The Order of Things* (Foucault 1970) is his most systematic attempt to investigate this.

[33]Foucault 1977a, part 1.

[34]Ibid., pp. 57–69.

[35]Ibid., part 2, section 1.

and contrived – and the associated lesson was a difficult and challenging one.[36] Ultimately, the successful replacement for the punitive regime founded on state violence and retribution was not the elaborate symbolic alternative, but the disciplinary system of the prison, which rapidly supplanted other approaches at the beginning of the nineteenth century.

The most distinctive feature of the prison, in comparison with other penal regimes, was that it aimed above all to make the criminal the subject of meticulous physical control, and observation.[37] The best illustration of the new system at work is Jeremy Bentham's model prison (which was never built, but whose architectural principles were embodied in a number of prisons and other institutional buildings) the 'Panopticon'.[38] The Panopticon consists of a ring of glass-walled cells, surrounding a central observation tower. Prisoners in the cells are constantly visible from the tower, silhouetted against the light. The tower itself is designed with baffles, so that the prisoners cannot see their overseers. Prisoners in the Panopticon are to be treated humanely – gone are the days of overwhelming state violence. Moreover, they are not to be the subjects of any sustained communication. No one will set out to directly communicate a lesson. Instead they will be reformed by a regime of unremitting surveillance. The Panopticon doesn't simply restrict the prisoners' movements. Nor does it simply combine physical control with observation. Rather, it is geared to ensuring that, in addition to being observed, the prisoner will know himself to be the subject of control and observation.[39] Thus while the Panopticon operates in part by means of an 'objectification' – prisoners are treated both as 'things' to be controlled, and as objects of knowledge – the end toward which this regime of surveillance serves as a means is that of inculcating the habit of self-surveillance in the prisoner. The aim is to procure, through the objectification of the prisoner, his self-objectification. Or, what amounts to the same thing, his 'subjectification', or *subjection*.[40]

The Panopticon runs on self-knowledge. Once the prisoners acquire the habit of self-surveillance the mechanism will, in principle, continue to operate without an observer in the tower. As with the confessional, the purpose of observation is not that of intelligence-gathering (to that end it would be more effective if the surveillance was covert). We are not dealing here with an example of knowledge which serves authority directly and straightforwardly as an instrument of power. Rather, the particular power-knowledge constellation exemplified by the Panopticon depends on the inculcation of disciplinary self-knowledge. Again, as with the confessional, the aim is to produce a body of knowledge which takes its effect in the subject of that knowledge. The self-knowledge that is key to the functioning of the Panopticon is intended to have a specific effect in the mind of the prisoner. It is addressed to

---

[36] Ibid., pp. 110–13.

[37] Ibid., pp. 120–31.

[38] Ibid., pp. 200–4.

[39] Ibid., p. 201.

[40] Ibid., pp. 201–3. See also Foucault 1981, p. 60.

him as subject, to his subjectivity, in that it aims to produce in him a constant and inescapable understanding of himself as the subject of both control and observation. The disciplinary power exerted within the Panopticon has its operative basis in the prisoners' self-conception. If we discount the prisoner-as-subject, we lose the crucial element in its operation.[41]

If the Panopticon effectively symbolises and explains the operation of disciplinary power, we need at the same time to look beyond the prison as an institution in order to appreciate its full reach. It is not only as the subject of physical self-surveillance that the prisoner is subject to disciplinary power. Alongside the prison, new techniques of knowledge-acquisition via 'examination' arose.[42] These served to make the penal subject an object of knowledge, as an individual, as never before. These individual case histories make new forms of criminological self-knowledge possible. The prisoner is henceforth aware of himself as the object of scientific knowledge, and, potentially, as the subject of scientific self-knowledge, in a manner that previously would not have been possible.

Ultimately then, Foucault's conception of disciplinary power is of a scientific/punitive complex, which unites the prison system and the nascent sciences of criminology, psychiatry etc., and operates by making the offender aware of himself as both an object of close control, and an object of close observation. The purpose of the scientific self-knowledge thereby made available is not the direct instrumental control of offenders, but their indirect control, since such control is mediated by their own forms of self-understanding. Scientific self-knowledge is here not a means of emancipation, but a means of (political and epistemic) subjection. Finally, the effects of disciplinary power are not felt at the level of conscious reflection, but in the 'habits of soul' that the combination of control and surveillance which is characteristic of disciplinary institutions serves to establish.[43] The same forms of control – systematic observation leading to systematic self-surveillance – are evident across a range of early nineteenth century institutions, from military camps to hospital wards.[44] Arguably the same forms are stronger than ever today, as the potential for 'subjectification' through scientific self-knowledge has continued to develop.

10.2   Subjection and Recognition

Foucault's analysis of disciplinary subjection displays many connections with the Sartrean and Heideggerian accounts of sadism and inauthenticity. There is, for example, the same concern with a form of instrumental stereotyping as a

---

[41] For further discussion of this point see Lucas 2002.

[42] Foucault 1977a, pp. 187–94.

[43] Ibid., p. 128.

[44] Ibid., pp. 170–7.

ubiquitous feature of human relations, and with debased and debasing forms of self-knowledge. The most striking differences between Foucault's analysis and those of his existentialist predecessors however are that Foucauldian disciplinary power is institutionalised, and effectively agentless. These differences are summed up in the image of the absent watcher in the central tower of the Panopticon.

Had Heidegger turned his attention to the Panopticon and related technologies of social control he would no doubt have emphasised the debasement of humanity within such institutions. Possible themes would have included the inauthenticity of the prisoners, who come to view themselves instrumentally, as cogs in a machine. He would perhaps also have been concerned with the inauthenticity of the overseer, who is likely to draw his own self-understanding from his role as an adjunct of such a technology. In certain contexts he might have expressed concern at the way in which such technologies appear to reduce human beings to mere resources – to objectify humanity by reducing it to 'standing reserve'.

Sartre might have regarded the Panopticon as a metaphor for the sadomasochistic dramas that are for him characteristic of human personal relationships. The overseer in the tower might serve as a symbol for pure being-for-itself, the prisoners as pure being-in-itself – subject, in their glass-walled cells, to the look of the other, in a particularly striking form.

Foucault, by contrast, emphasises the fact that the Panopticon is able to operate – for much of the time at least – independent of any controlling agent. Just as the confessional has its power effect on the confessee irrespective of whether anyone is listening, so the panoptic system has its power effect on the prisoners directly, by affecting their self-conceptions, without the need for an overseer in the tower. For the prisoners to be subjectified in this way it is no doubt necessary that, at some stage, they have grounds for thinking they may be observed. But such grounds can exist without their actually being observed, at any particular point. And it seems that they could exist, in principle at least, without there ever being an overseer in the tower.

Panoptic systems have this quality because the essential moment in subjection is the moment of *recognition*. This is not, needless to say, recognition in the Hegelian sense – involving an encounter with someone who can provide me with the external perspective I need if I am to achieve self-actualisation. But it is not entirely unrelated. It is much closer though to the recognition that Sartre regards as an essential moment in shame, and in sadism. Foucault however is concerned with mechanisms that operate at the level of historical processes, social institutions, and scientific disciplines. We are subjectified by being brought, through a process of (real or imagined) surveillance, to particular forms of self-understanding. I become the subject of disciplinary power/knowledge when I recognise myself in the object it describes.

The Foucauldian model is therefore able to address any concerns that might arise in connection with the voluntarism and the generality of the Sartrean model. We are directed to structural factors, and features of historical institutions, in order to understand how debasing forms of self-knowledge are generated and perpetuated. The relative agentlessness of the Foucauldian model also has implications for the type of instrumentalisation it should be understood to involve. In the

Sartrean account the implication was that the *content* of the victim's self-conception will reflect some particular purpose that the sadist has in mind. But, of course, if the real purpose is simply sadism, the precise content of the victim's self-conception is of secondary importance, provided it achieves the desired effect. The Foucauldian model reflects this, concentrating not on the instrumental derivation of the self-conceptions that disciplinary power/knowledge makes available, but on the instrumental value of their adoption. Foucault's emphasis, then, is less on instrumentalised self-understanding, than on self-understanding as an instrument. And, of course, there need not be, at any given time, anyone except the victim actively wielding the instrument.

The account of panoptic power does not exhaust Foucault's analysis of disciplinary power/knowledge, but it does illustrate clearly its main features, and its modus operandi. By contrast with Sartre's explicitly conflictual model of human relationships, Foucault's account of the objectifying processes at work within disciplinary institutions portrays them as operating only semi-self-consciously. By comparison with the excesses of the old monarchical punitive system, the disciplinary power/knowledge system might be considered to be a relatively unspectacular manifestation of human sadism. It would be characterised above all by its habituality, its banality. But that would not make it insignificant. It was not Foucault's aim to analyse the more glaring instances of human cruelty. Rather, he sought to identify more subtle forms, belonging to what might be thought to be relatively enlightened social institutions.

10.3   Implications

The views we have considered over the last three chapters supply us with the materials we need for a coherent, rich and defensible sceptical essentialist account of the origins of objectification/subjectification, and associated interpretive moral wrongs.

Sartre takes up the Heideggerian sceptical-essentialist conception of inauthenticity as self-instrumentalisation/objectification, but attaches it to a philosophy of human relations in which the other figures as a genuine other-like-me. This sets the scene for the 'game of mirrors', in which objectification can be doubly reflexive – not simply a matter of self-interpretation, but self-interpretation in light of an experience of the other as a self-interpreting being, and a source of alternative interpretations. At the social level, the relatively organic-seeming model advanced by Heidegger is replaced by a conflictual model of human relations, in which subject-objects struggle with themselves and others to get free of the instrumental and instrumentally-derived stereotypes that press in upon them.

Foucault, while distancing himself from the phenomenological notion of the meaning-constituting subject, develops certain recognisably Sartrean themes, albeit with much greater breadth of social and historical reference. He examines how the battle of self and other has been played out historically, through linked political and scientific struggles, and revolutions. His analysis also highlights the political utility

of doubly reflexive objectification. The prisoners in the Panopticon, the citizens of a disciplinary society, are put under surveillance not for the direct purpose of information gathering but because being subjected to intense observation will have its effect on their self-conception, and subsequently on their behaviour. They are 'objectified' mendaciously. They are treated as objects, and yet the entire procedure is premised on the knowledge that they are (potentially at least) self-knowing subjects.

Foucault's analysis of (substantially) agentless disciplinary power/knowledge is thus highly relevant to anyone concerned with the ethics of respect for self-interpreting agents, in the context of professional-client interactions – particularly in contexts that involve the deployment of authoritative but ambiguous bodies of knowledge. His historico-philosophical studies are an important supplement to the socially and politically restricted focus of Sartre's account of concrete relations with others .

References

de Beauvoir, Simone. 1952. *The second sex* (trans: Parshley, H.M.). New York, NY: Alfred A. Knopf.

Foucault, Michel. 1970. *The order of things.* London: Tavistock.

Foucault, Michel. 1977a. *Discipline and punish* (trans: Sheridan, A.). London: Penguin.

Foucault, Michel. 1977b. Nietzsche, genealogy, history. In *Language, counter-memory, practice*, ed. Donald F. Bouchard, 139–164. Ithaca, NY: Cornell University Press.

Foucault, Michel. 1980a. *Power/knowledge: Selected interviews and other writings, 1972–1977.* Hemel Hempstead: Harvester Wheatsheaf.

Foucault, Michel. 1980b. Truth and power. In *Power/knowledge*, ed. Colin Gordon, 109–133. Hemel Hempstead: Harvester Press.

Foucault, Michel. 1981. *The will to knowledge: The history of sexuality*, Vol. 1 (trans: Hurley, R.). London: Penguin.

Foucault, Michel. 1988a. Critical theory/intellectual history. In *Politics, philosophy, culture: Interviews and other writings 1977–1984*, ed. Lawrence D. Kritzman, 17–46. London: Routledge.

Foucault, Michel, 1988b. The return of morality. In *Politics, philosophy, culture: Interviews and other writings 1977–1984*, ed. Lawrence D. Kritzman, 242–254. London: Routledge.

Foucault, Michel. 1989a. *The archaeology of knowledge* (trans: Sheridan, A.M.). London: Routledge.

Foucault, Michel. 1989b. *The birth of the clinic* (trans: Sheridan, A.M.). London: Routledge.

Heidegger, Martin. 1962. *Being and time* (trans: Macquarrie, J. and Robinson, E.). Oxford: Blackwell.

Lucas, Peter. 2002. Mind-forg'd manacles and habits of the soul: Foucault's debt to Heidegger. *Philosophy of the Social Sciences* 32(2):310–328.

Nietzsche, Friedrich. 1966. Beyond good and evil. In *Basic writings of Nietzsche*, ed. Walter Kaufman, 179–435. New York, NY: Random House.

Nietzsche, Friedrich. 1974. *The gay science* (trans: Kaufman, W.). New York, NY: Vintage Books.

Sartre, Jean-Paul. 1989. *Being and nothingness* (trans: Barnes, H.). London: Routledge.

Chapter 11
Honesty

Abstract The post-Hegelian ethic of self-interpretation developed in Chapters 8, 9, and 10 might still be regarded as normatively weak, insofar as, in giving up essentialism, we seem to give up any appeal to the notion of a distorted self-conception. However, Nietzsche's thought supplies what we need to give useful sense to the notion of a distorted self-conception, while remaining sceptical of essentialism. Nietzsche's emphasis on self-overcoming is grounded in a fundamental appeal to *honesty* in self-interpretation. Similarly, the sceptical essentialist can re-interpret references to distorted self-conceptions as references to self-conceptions that would, if adopted, be adopted dishonestly. Moreover, since I am not self-sufficient in my self-interpretations, such honesty would have to have a collective basis. Collective honesty would be a suitable foundation for a sceptical essentialist ethic of self-interpretation. The chapter closes with some reflections on the implications of this view for professional ethics.

11.1 Honesty Versus Self-Creation

> Whom do you call bad? – *Those who always want to put to shame*
> What do you consider most humane? – *To spare someone shame*
> What is the seal of liberation? – *No longer being ashamed in front of oneself*
> *Nietzsche, The Gay Science.*

I first raised ethical concerns about the ideological stereotyping of ontological persons in Chapter 4, and went on to consider how, particularly in conjunction with distributive injustice and instrumentalisation, such stereotyping is linked to interpretive moral wrongs. A concern with stereotypes that are straightforwardly false was set on one side. This was not because such stereotypes are not ethically and politically significant, for they are. The motive was simply that such stereotyping is quite evidently of ethical significance, and is, consequently, relatively straightforward to challenge. Where an apparently false stereotype proves to be relatively durable, and resistant to challenge, this will usually be because it is closely connected with the more insidious forms that I have referred to as 'ideological' stereotypes. These are categorisation/generalisation complexes that cannot be shown to be straightforwardly false at either the general or the particular level, because they serve to

structure our thought.[1] Such stereotypes may be associated with large sets of true beliefs concerning relevant groups, and the difficulty of challenging them may stem partly from an automatic and unthinking deference toward 'the truth': if a belief, or cluster of beliefs, is true, it might be argued, what possible objection can there be to its promulgation?

The proper strategy when dealing with an ideological stereotype is not necessarily to challenge the truth of the various generalisations it embodies, but to challenge the legitimacy of the categorisation-generalisation complex that it embodies. That a given categorisation-generalisation complex contains some truth does not entail that it is legitimate. Highly arbitrary classifications, with little or no predictive or explanatory power (e.g. 'people with the same number of letters in their forename and family name') could be used as the basis for true generalisations, but would lack all legitimacy as classifications (since the generalisations in question would be true merely accidentally). Moreover, some categorisation-generalisation complexes which (perhaps in conjunction with some more wide-ranging socio-economic theory) do yield predictive and explanatory power, fail to qualify as legitimate bases on which to assess or judge persons, because the human realities they reflect are the legacy of a regrettable history of social injustice. In such cases the generalisations in question would, where true, be true only as a matter of historical contingency. The principles underlying anti-discrimination policies and legislation constitute an implicit acknowledgement that we are entitled, particularly in matters that significantly affect our economic well-being, and legal or civil freedoms, to be assessed according to standards that reflect what we would have been in the absence of regrettable socio-historical factors.

When ideological stereotyping is combined with instrumentalisation we get objectification. Perhaps the most obvious example of objectification is where the instrumentalisation of a group is accompanied by the implicit or explicit claim that the members of the group are, in effect, 'natural' slaves.[2] This sort of assessment may be linked to a false stereotype, or it may be linked to an ideological stereotype, associated with true generalisations, which, however, only reflect the damage that a history of instrumentalisation has caused. More subtle forms of objectification include the treatment of persons as specimens or models in e.g. biomedical scientific research, and the sexual fetishisation of women in pornography, advertising etc. In all of these cases identifiable ideological stereotypes may be linked to some form of instrumentalisation.

It may be tempting to focus exclusively on the instrumentalising element in objectification, since the reduction of human beings to the status of a means has long been regarded as ethically unacceptable. But such an approach overlooks the distinctive wrongs associated with stereotyping itself. To judge someone on the basis

---

[1]Much as do the Foucauldian *episteme* (Foucault 1970) and the Kuhnian *paradigm* (Kuhn 1962) in the natural and human sciences.

[2]As in Aristotle 1962, book 1, chapter 13.

of an ideological stereotype is to fail to show proper respect for their human dignity, whether or not any further harm or wrong results. Indeed using someone as a means while stereotyping them looks ethically unacceptable even in cases in which they are arguably not reduced to the status of a mere means. Marx's analysis of the commodification of the worker under capitalism provides a convenient illustration of this latter point.[3] The distinction between slavery and waged labour is often used to illustrate the difference between using someone as a means, and using them as a mere means. But where the waged labourer is simultaneously the object of an economic theory that treats her labour power as a commodity, it seems reasonable to say that a form of reduction to the status of a mere means has occurred, whether or not the worker is paid for her labour. Where instrumentalism is accompanied by stereotyping then, we cannot safely assume that the traditional version of the distinction between serving as a means, and being reduced to a mere means, holds good.

Marx's analysis also introduces the important point that the effects of instrumentalisation on the self-conception of the instrumentalised party are ethically and politically significant. If there is a wrong associated with understanding another as a commodity, independently of whether any equally dramatic 'reduction' is evident in the terms of her employment, then there is surely also a wrong involved in inducing the worker to understand herself as a commodity (reducing her to the status of a means in her own eyes). Thus the associated wrongs seem to be in large part a function of the reflective capacities of potential victims. The consequences for the victim's dignity are not simply a matter of the way she is treated, and conceived of by others, but also a matter of how, consequently, she comes to conceive of herself.

In order to further explore the ethical significance of offences that affect the victim at the level of her self-understanding, and to detach our analysis from the essentialistic presuppositions of Hegel and Marx, we surveyed a range of post-Hegelian views on the ethics and politics of recognition and self-knowledge. In common with Hegel, these post-Hegelian thinkers are centrally concerned with the ethics of recognition and self-interpretation. But they differ from Hegel on the key matter of essentialism. Where Hegel understands the cognitive and moral development of humanity as a teleological process, tending to a definite end, Heidegger and Sartre reject Hegel's teleological essentialism in favour of an existentialist view, according to which the only legitimate sense that can be given to the idea of actualising our human essence is that involved in the struggle for authenticity, or the avoidance of bad faith: we actualise our essence insofar as we remain resolutely open to our own unique possibilities, and steadfastly refuse the consolations of a teleological philosophy of human nature. Foucault's rejection of essentialism and teleological thinking is more straightforward, and dispenses with any lingering air of paradoxicality. But it is equally uncompromising.[4]

---

[3]Marx 1967, p. 289.
[4]See e.g. Foucault 1977.

Although Hegel accepts that self-knowledge may be more or less adequate (and this possibility is central to his conception of self-actualisation), he seems optimistic that the process of coming to understand ourselves as objects will prove to be a temporary, if necessary, step on the ladder to genuine self-knowledge. The idea that our progress might be deliberately arrested at that stage, and our objectified self-understanding might become a point of extreme vulnerability, is not one he explicitly concerns himself with. By contrast, our trio of later thinkers are united by a distinct pessimism concerning the benefits of self-knowledge – particularly (in the case of Foucault and Heidegger at least) a pessimism concerning the benefits of scientific self-knowledge. It would seem that the origins of this pessimism lie in their combining a Hegelian conception of the importance of recognition and self-understanding, with a view of the origins and nature of knowledge derived in large part from Nietzsche.

Notoriously, Nietzsche views knowledge, including scientific knowledge, as an expression of will to power.[5] What we take to be truths are really more or less useful errors.[6] There is, in Nietzsche's view, no possibility of grounding knowledge in some extra-human realm – either a metaphysical beyond (as Plato imagined), or in a teleological nature (as Aristotle seems to have believed, and as modern essentialists such as Hegel and Marx believed). What we find instead are beliefs and belief systems, succeeding each other in history, in a manner that is ultimately directionless. These beliefs are the effects and instruments of power. They are means by which groups of humans have exerted power over other humans, and over the natural world. And the aura of truth, objectivity and authority that clings to them is, for Nietzsche, an effect of the disparities of power that lie at their source.[7] Power creates and utilises knowledge, and gives the appearance of neutrality and authority to what is in fact always an expression of, and a means of furthering, particular human interests.

The influence of Nietzsche's conception of knowledge on Heidegger, Sartre and Foucault is readily apparent. All three think of self-knowledge as intimately connected with power. This connection is most obvious in the work of Foucault, and runs as a constant thread throughout his writings. But it is also present in the thought of Sartre, as our examination of his view on concrete relations with others revealed (Section 9.4). For Sartre, the forms of power that are of most concern are those in play in personal relationships. But there seems to be nothing to prevent the sadomasochistic patterns he describes being played out at a more general level, between social groups. Heidegger, as we saw in Chapter 8, is concerned with inauthentic self-understanding primarily at the level of the individual. The *politics* of self-knowledge, either at the micro level, in personal relationships, or the macro level, between social groups, do not really seem to concern him. And yet the forms of inauthentic self-understanding he discusses are intimately bound up with power

---

[5]Nietzsche 1954b, part 2, 'On Self-Overcoming'.

[6]Nietzsche 1974, sections 110–4.

[7]Ibid., section 110.

relations, in that they are typically drawn from human practical activity, geared to transforming our world.[8]

It is clear that this combination of a Hegelian concern with the ethics of recognition and self-interpretation, and a Nietzschean view of the power/knowledge connection, will tend to generate an anti-teleological and highly critical view of the epistemic, social and ethical consequences of supposed advances in scientific self-knowledge and self-understanding. But for all the critical sound and fury such approaches are apt to inspire, it may still reasonably be doubted whether they can generate sufficiently strong normative conclusions. Consider the basis on which any such approach might challenge ideological stereotyping. Challenging an ideological stereotype involves calling into question a particular conception of human nature, or of the nature of some specific group of humans. In order to do this, it may not be necessary, in practice, to make an overt appeal to a conception of a human essence. But the theoretical problem remains: on what grounds can an approach that rejects essentialism engage in such criticism?

Earlier (Section 4.3) I characterised ideological stereotypes as conceptions of the nature of groups of ontological persons that embody distortions, and are distorting, when taken up into the self-understanding of the victim. An essentialistic ethic of recognition, such as Marx's Hegelian critique of commodification, is equipped to reject ideological stereotyping because it can appeal to a substantive conception of human nature, by comparison with which the commodified self-conception of the industrial worker appears distorted, and distorting. But the Nietzsche-influenced views we have examined cannot consistently adopt this strategy. The problem is not that they are not equipped to reject a given ideological stereotype as, when taken up in an individual's self-understanding, an example of inauthenticity or bad faith. The problem is that this criticism appears to apply equally to any substantive view of the self whatever. All self-conceptions seem to be equally inauthentic. More accurately: authenticity and inauthenticity are not tied to any particular self-conception, except at the most formal level (at which authenticity is a matter of embracing one's unique possibilities, inauthenticity a matter of seeking to evade them). Thus they are more like modes of relation to a given self-conception. It seems conceivable then that virtually any self-conception might be rehabilitated, on the condition that it was held or adopted in the 'right' way, or rejected, on the basis that it was held or adopted in the wrong way. At this point then, our Nietzsche-inspired views seem in danger of losing all normative purchase.

There is however more to be said in favour of such a view than this. My original motive for turning to these thinkers was that their views lack the problematic essentialism of Hegel's account of recognition and self-actualisation. Although none of these thinkers was a political liberal, their views are, in this respect at least, likely to be generally more congenial to a generally liberal political culture. Richard Rorty has indicated how this kind of philosophical approach can be squared with liberal

---

[8]Heidegger 1962, sections 25–7.

values, while retaining a degree of normative appeal.[9] Rorty sees the origins of such views in the post-enlightenment idea that 'truth is made rather than found'.[10] To the extent that the claims of Marx and other radicals concerning the alienation of industrial workers represent a social advance, they do so not because they expose deeply buried truths that classical economics strives to keep hidden, but simply because they sketch better (because more humane) ways of organising human society.

For Rorty, all self-interpretation is really nothing but a form of self-creation; and philosophical approaches to the ethics of self-interpretation can either be candid about this fact, or can attempt to suppress it. If they take the former course they will be openly *utopian*.[11] Their characteristic rhetorical stance will be to say, in effect: 'we cannot say that the way we live now is wrong or distorted/distorting in any ultimate sense, but we can imagine ways to live that are undeniably better. . .'. Such utopian visions are both the most straightforward, and, in Rorty's view, historically the most effective ways to motivate genuine political change.[12] Moreover, they are not only compatible with political liberalism, they are vital for the health of a liberal political culture – since the possibility of engaging in imaginative experiments in living is an important component of political liberty.

Nonetheless, for Rorty, it is important that this process of envisioning new human possibilities should remain utopian, rather than *radical*. By a 'radical' strategy for change he has in mind the sort of political approach (which often presents itself as a scientific approach to solving the problems of human history), that suggests that an important error has been made deep down and long ago, which it is the task of careful philosophical and scientific analysis to diagnose and address.[13] Such a view will tend to contrast what human beings really are (and, relatedly, how human society should really work), with what has historically been the case. What is in fact going on, in Rorty's view, when anyone does this, is that they are engaging in imaginative utopian self-creation. They are not really painstakingly pinning down what human beings are, they are imagining ways in which they could be different. But, according to Rorty, radical strategies for change suppress the creative aspect of what they are doing, often so as to lay claim to a spurious scientific legitimacy.[14]

Apparently to assuage liberal concerns that utopianism tends to lead to the worst kind of totalitarian excesses, Rorty adds that good liberals should be 'ironic' about their public commitments. Utopian experiments in living are best kept for the private rather than the public arena; and, as projects of self-creation, they needn't, and ideally shouldn't, inspire the sort of revolutionary zeal associated with radical strategies, which are structurally committed to the idea that moral and political progress

---

[9]See Rorty 1989.

[10]Ibid., p. 3.

[11]Ibid., Introduction. See also Rorty 1991.

[12]Rorty 1989, Chapters 7–9.

[13]See Rorty 1991.

[14]Ibid., p. 7.

brings us closer to the truth.[15] (In acknowledgement of liberal qualms about utopianism, Rorty would no doubt also point out that much of what Popper refers to as 'utopian' political theorising is actually radical theorising, and not properly utopian at all.) In recommending that political liberals should preserve an ironic separation between their public commitments and their private projects of self-creation, Rorty is not suggesting that they should not take the latter projects seriously. His concern is simply to preserve the sense that such creative visions of what human beings could be belong to projects of 'self-overcoming'. We should never make the mistake of essentialising them.

However, in its rejection of radicalism, this approach still seems normatively weak. Earlier I presented ideological stereotypes as distorted and distorting categorisation-generalisation complexes. A Rortian liberal approach would tend to suggest that such stereotypes may be limiting. But, for Rorty, only an essentialist radical could claim that they are *distorting*. We cannot, within Rorty's model, say that such stereotypes get humanity *wrong*, we can only say that certain better bases for self-interpretation are possible, and are to be recommended.

One possible recourse at this point would be to protest at the instrumentalising quality of any particular self-interpretation. If we subscribe to the kind of essentialism that says that, whatever else we may be, humans are not properly conceived as instruments, then it will be possible to reject some self-conceptions on the basis that they are instrumentalising, without broaching the question whether they are, in any more specific sense, distorting. Unfortunately though, the Nietzsche-influenced views we have considered seem bound to consider any self-conception to be instrumental, insofar as they think of knowledge in general as an effect and instrument of power. What possible self-conception would not be instrumentalising? And how, consequently, are we to distinguish acceptable from unacceptable self-interpretations?

Again, then, we seem to be thrown back on an unsatisfactory choice – though this time it is not the choice between essentialism and a view that appears to lack all normative purchase, but a choice between essentialism and the rather evanescent ethical appeal of exercises in 'ironic' self-creation. Is there any way, then, to develop an ethic of recognition and self-interpretation that has more power to combat ideological stereotyping, but at the same time avoids falling into a problematic form of essentialism?

Initially, it is important to remind ourselves that none of the views we have been considering is debarred, in principle, from appealing to an idea of human dignity – of the type discussed at the close of Chapter 6 – in opposing ideological stereotyping, and the interpretive moral wrongs associated with it. It is possible to claim that a particular stereotype is distorting, as being (1) untrue, or (2) true only accidentally, or (3) true only as a matter of historical contingency, without necessarily being drawn into any deeper discussion of what *grounds* these claims about distortion, and respecting human dignity. Kantians, Hegelians, Foucauldians and Rortians can all in

---

[15]Rorty 1989, pp. 73–8.

practice make common cause in claiming that there is such a thing as human dignity, and that dignity is compromised both when people are treated as mere instruments or resources, and when they are conceived of in line with some ideological stereotype, without having to air their differences on the question of what grounds such claims. (Whether, that is, they are to be grounded with reference to some essential human nature, or (e.g.) simply with reference to some more attractive imagined alternative.)

That said, in a theoretical context, we are also going to want to know whether, in light of the findings of the post-Hegelian tradition we have been tracing, there is any alternative to Rorty's utopian strategy, once essentialism is abandoned. I believe that there is. To appreciate this possibility it is helpful to consider again the example of Nietzsche. It is a surprising yet undeniable fact that despite his unrelenting attacks on traditional views of morality, Nietzsche was a deeply moral thinker. Indeed, it could be argued that one fails to grasp the true character of his philosophy unless one appreciates that these two facets of his thinking are intimately connected. The connection lies in his uncompromising commitment to *honesty*. Nietzsche is determined to root out all elements of self-delusion. His rejection of traditional morality, and of traditional views of knowledge, is not undertaken purely for its own sake, but partly as a consequence of this uncompromising demand for honesty.[16]

In this his position has much in common with ancient scepticism. The ancient sceptics did not reject the claim to know on the basis of a simple hostility toward knowledge. Rather, they refused to believe anything for which sufficiently good reasons to believe could not be given (observing at the same time that the absence of sufficiently good reasons to believe seemed to be quite general).[17] Nietzsche too is sceptical about traditional views of knowledge and morality. It is not however from sheer perversity that he refuses to believe. Rather it is from a determination not to indulge his desire to believe, if it should turn out that adequate reasons to believe cannot be found. It follows that one misses a key aspect of Nietzsche's thought if one sums up his view, as Rorty's presentation of liberal ironism and utopian self-creation as processes of 'self-overcoming' suggest it might be summed up, as a straightforward application of the insight that 'truth is made and not found' to the process of self-interpretation. Nietzsche's conviction is that the drive to self-overcoming originates from ruthless honesty.[18] If there is to be any self-creation, it must be honest self-creation. The point is not simply that we must invent better ways of being

---

[16]Nietzsche 1974, section 319; 1966, sections 227, 229, 230.

[17]Such, in broad outline at least, seems to have been the view of Pyrrho of Elis, for example.

[18]Nietzsche 1974, sections 110–114, 319. It is true that in the 1873 fragment 'On Truth and Lie in an Extra-Moral Sense' (Nietzsche 1954a, pp. 46–7) he expresses doubts about the explanatory power of the appeal to honesty, faced with the variety of individual acts to which the term might be applied. But the problems afflicting the term 'honesty' prove here to be the same as those that afflict all general concepts: we treat unequals as if they were equal. While this problem clearly has special relevance to the concept of honesty, once acknowledged it gives us no more reason to abandon that particular general term than to abandon any other general term. Thus the point does not altogether undermine the idea that individual acts, views and opinions can be characterised as more or less honest.

human then, but that we must be merciless in our self-interpretation – rooting out ideology wherever we find it, even if this means we find there is nothing at all we can believe in. At least then our unbelief will be the manifestation of a key virtue, namely honesty.

This fundamental commitment to honesty gives vastly more depth to the project of 'self-overcoming'. It also suggests a plausible sense for the existentialist notion that to be human is to be a being whose being must always remain an issue. Such a being does not simply engage in frivolous self-creation, as if the insight that Hegelian essentialism is untenable licensed any and every experiment in living. The thrust of the insight is that in refusing any particular self-interpretation we display a capacity that is of key importance to being human: the capacity to refuse self-conceptions not on the basis of taste, but on the basis of (moral and epistemological) principle.[19]

The characteristically Nietzschean virtue of honesty will serve as a reliable basis for a sceptical essentialist ethic of self-interpretation. Rejecting the idea that our nature is simply (potentially) discoverable does not commit us to embracing the alternative view that all we can say of competing self-conceptions is that they are (better or worse) inventions. A given self-conception can still coherently be rejected on the basis that to adopt it would be dishonest. Thus the notion of a distorted self-conception can be glossed as a self-conception that would, if adopted, be adopted dishonestly.

In all of this, and in line with the paradoxical existentialist doctrine that our essence is to interrogate our own essence, there is a fundamental self-conception at work, and so, in a sense, a form of essentialism. But it is an essentialism of the most minimal kind. It is a sceptical essentialism, that understands the fundamental human virtue to be that of honesty, and regards this virtue as a power capable of dismantling any and every moral and epistemological system. Nietzsche's attachment to honesty is not a last vestige of a discredited essentialism, maintained against his own better judgement. It is the practical basis on which he sets out to demolish untenable systems of thought.[20]

Such moral and epistemological scepticism is not therefore total scepticism – at least not in its moral dimension. The refusal to believe without good reason is itself a moral stance – albeit a minimal one and (in my view at least) a relatively uncontroversial and intrinsically appealing one. *This* critical spirit is, in an absolutely minimal sense, a radical, essentialist spirit. It makes its moral stand on a narrow but crucially important ledge: the human capacity to refuse, in spite of everything, to lie to ourselves – most of all, to refuse to lie to ourselves about ourselves.

---

[19] Note that this stance gives us a basis for rejecting truths, as well as falsehoods. Referring back to my earlier discussion of ideological stereotypes (Section 4.3), and the fact that particular categorisation-generalisation complexes may comprise a number of general truths, even while the complex itself lacks all legitimacy: Nietzschean honesty gives us a basis on which to reject such complexes, however many truths they may embody.

[20] Nietzsche 1974, section 319.

11.2 Sceptical Essentialism and Collective Honesty

The exercise of the kind of sceptical essentialist honesty outlined above may look to be a lonely sort of activity. While possibly informing a project of private perfection, how could it inform a system of public obligations? How could it become a collective ethic? Presumably, the individual commitment to honesty espoused by Nietzsche and Foucault must become a collective commitment. But what would motivate such a move? *Why* should the virtue of honesty express itself collectively?

The answer lies in the notion of objectivity, and the feature of objectivity that we have already noted: that for an understanding of objectivity I am dependent on others.[21] The commitment to honesty characteristic of Nietzschean scepticism is a commitment to believe only on the basis of adequate reasons (coupled with the observation that existing moral and epistemological principles seem to be adopted without adequate reasons). But such a commitment would be meaningless without some notion of objectivity in the background. I cannot possibly specify what it would be to believe on the basis of adequate reasons without invoking an idea of what would be accepted by an ideal community of enquirers, or at least by some epistemological saint, situated as I am situated. Such a stance is not then quite as heroically individualistic as it may at first appear. The very idea of having reason to believe involves an implicit reference to another, or others, and an implicit dependence on others, since it involves a notion of what *anyone* in our position ought to believe.

Of course, it might be responded that what is a reason to believe for one person may fail to be a reason for another. But if this is sometimes the case (and no doubt it is) then this is at the same time a contingent matter. Either the claim is that what is a good reason for me to believe is necessarily a good reason for anyone similarly situated to believe, or it is that a reason may be a reason for one person and one person only, however situated. In the former case it amounts to the claim that what is a reason for me will be a reason for anyone in my position, albeit with the rider that, as a matter of fact, no-one else is quite in my position. In this form the claim implicitly acknowledges that the idea of a reason to believe is a fundamentally social one. Or, in the latter case, the claim looks like the outcome of a failure to grasp what a reason to believe is. If you say that what is a reason for you need not be a reason for anyone else, however situated, then this suggests you have not understood what distinguishes a reason from a motive.

The point here has a distinctly Hegelian character: all rational critical thought presupposes the idea of intersubjectively valid principles, and consequently presupposes a form of potentially universal self-consciousness. Honesty without objectivity is nothing, and objectivity is a fundamentally social notion. (Though it does not of course follow that any social group need ever be in a position to stipulate what counts as objectively true.) Evidently then, Nietzschean honesty, however individualistic seeming, cannot maintain itself without some implicit reference to

---

[21]See Chapters 7, 8, and 9, above.

the social. The virtue of honesty must have a collective dimension. It is not that we cannot be honest outside of some actual community of enquirers – Nietzsche repeatedly emphasises how rare and remarkable the thoroughly honest individual would be – rather, the point is that such honesty involves a commitment to a conception of collective honesty. I am not wholly self-sufficient when it comes to my capacity for honesty. I may not rely on the de facto support of others, but I do rely on the idea of a possible community of enquirers who share a commitment to believe only on the basis of good reasons. It is on the basis of such collective honesty that I may reject a dominant ideology, as being fundamentally dishonest, without in doing so abandoning my commitment to objectivity and the possibility of mutually recognising ideology as ideology.

Collective honesty thus understood could form the moral basis of a sceptical essentialist opposition to ideological stereotyping, objectification, and other interpretive moral wrongs. The basic distinguishing feature of such a view would be that honesty with respect to our self-interpretations is a key virtue for beings like ourselves, where honesty is understood to involve a commitment to objectivity – the idea of believing only on the basis of (in principle) intersubjectively valid reasons – and where such objectivity is thought of as founded in ideal collectivity.

My argument to this point has claimed to reveal a fundamental, if minimal, essentialism at the basis of Nietzsche-inspired opposition to stereotyping and objectification. We noted earlier however that our liberal political culture tends to be hostile to all forms of essentialism. Does it follow, then, that such a view is ultimately incompatible with liberalism? I have pointed out (Chapter 7) that it seems unlikely that any liberal view could incorporate the type of radical critique I am advocating. (And Rorty's re-interpretation of self-overcoming as a project of self-creation provides further evidence of this.) Nevertheless, it should be pointed out that the approach has dispensed with many of the aspects of essentialism that liberals find offensive. As previously discussed, liberals treat Hegelian essentialism with great suspicion. The basis of their objections is the observation that Hegelian self-actualisation implies the existence of a general moral project, on which humanity as a whole is engaged. This Hegelian view is held to establish a contestable moral ideal as if it were a natural goal for human beings.[22] Popper's most virulent criticisms of Hegelianism concern the means by which such a moral ideal might be pursued in practice. But it is also clear that he considers Hegelianism to be necessarily incompatible with a rational, scientific, approach to ethics and politics. Liberals, it is supposed, are not committed to any unified moral project, and the attempt to enforce such a commitment would be at best counter-productive, at worst a totalitarian nightmare.[23] A rational approach to ethics and politics must, on Popper's view, be based on trial and error, rather than speculative philosophising.

---

[22]Wood 1990, pp. 33–5.
[23]Popper 1945, chapter 12.

But even in saying this, Popper reveals his own allegiance to a particular moral project: the project of rationally pursuing an objectively well-founded understanding of ourselves, and our place in the world. It may be that Popper is correct that this project is best served by trial and error experimentation, rather than by centralised planning. But the project itself is one in which we are all assumed to have an investment – and an investment that is not simply founded on pragmatism. Thus Popper's own criticisms of wholesale social engineering tend to suggest that the virtue of collective honesty, and the idea of a minimal essentialism founded on it, may not be entirely alien to his own liberal views. It is partly on the basis of an unconscious subscription to a view of collective honesty as a fundamental virtue that Popper's anti-essentialist stance has liberal appeal.

Sceptical essentialism founded on collective honesty ought then to have some appeal to political liberals. But would such essentialism have sufficient normative bite – more than might be possessed by Rorty's liberal ironism, for example? I believe that it would. The straightforward essentialism of Hegel and Marx implies a view of ideological stereotypes as distorted and distorting views of what it is to be a member of such and such a group. *This* idea of distortion must be founded on straightforward essentialism, because only the claim to have achieved some positive grasp of the human essence could back up the claim that a given self-conception is positively *distorted*. (Just as we can only sustain the claim that a given sketch is a caricature on the basis of having reliable access to the original subject of the sketch.) But the sceptical essentialism founded on collective honesty that I am advocating would not interpret stereotyping in this way. On such a view the proper gloss on the idea of 'distortion' would be that a given self-conception is distorting if it is dishonest. To subscribe to a stereotyped view of oneself or others would be to fail to manifest the virtue of honesty. It would be to believe on the basis of insufficient reasons, in an area where it is particularly important that we do not do so.

A sceptical essentialism founded on the virtue of (collective) honesty thus holds the resources we need for rejecting ideological stereotyping. We needn't simply devote our energies to utopian dreaming and self-creation – imagining better ways to be (though it is no doubt important to spend some of our time doing this). We can also devote our energies to radical criticism of existing ideological stereotypes, bringing out the way in which subscribing to them (even where the generalisations they involve are true) involves believing on the basis of notably inadequate reasons.

Can the virtue of honesty bear the weight I am asking it to carry here? If the central concern is to find some normative basis on which to challenge ideological stereotyping, it could be argued that honesty is an insufficiently powerful weapon against ideology – since it is possible to be honest, in the sense of being sincere, whilst still remaining in the grip of ideology. But is this really a credible worry? No doubt there is a restricted sense of honesty in which it amounts to nothing more than sincerity. But where (for example) we ask someone for an honest assessment of the potential success of an enterprise, though we would expect a sincere answer, we would not be content with an answer that was merely sincere. At the very least, we would expect the individual in question to have exercised due diligence in investigating the question on which she is being invited to pronounce. Simply

telling us what she thinks would not be sufficient for an honest assessment, except in contexts in which it is taken as read that she is already sufficiently familiar with the case. In cases like this, honesty does not consist simply in not lying or consciously deceiving. Rather, it involves responsible truth-telling, where possible. Thus while it might be possible to give an opinion that is both sincere and thoroughly ideological, it is not, I would contend, possible to give an opinion that is both genuinely honest and thoroughly ideological. I have already emphasised how established anti-discrimination safeguards require us to take active steps to avoid ideological stereotyping in specified circumstances. It would not be acceptable, in any such case, to protest that since one was sincere in applying one's ideological stereotype, one could be said to have done so honestly. While there are undoubtedly other values at stake in such cases, in addition to that of honesty, it seems an entirely appropriate response to point out that there would be a failure of honesty involved in proceeding on the basis of such a stereotype, however sincerely it might have been done.

It is true that ideology and honesty do not seem to be entirely mutually exclusive. But it should be borne in mind that, for someone committed to an ideal of honesty that goes beyond mere sincerity, it is not necessary that they should henceforth evade ideology on all points. If Nietzsche is to serve as our exemplar of this form of honesty (as Freud's observation that he displayed a more penetrating self-knowledge than any man who had ever lived suggests he might), then in arguing that, on balance, he lived up to that ideal, we can at the same time afford to concede that (as abundant evidence scattered throughout his writings attests), he certainly fell into the grip of ideology on some points. Nevertheless, when Nietzsche falls into the grip of ideology he evidently fails to live up to the standard of honesty he sets himself – it is a failure by his own lights. It is a commitment to honesty in this sense (if not necessarily to this degree) that is needed in order to put the sort of opposition to ideological stereotyping I am advocating into practice. Perhaps nobody ever evades ideology completely. However, a signal virtue of the sceptical essentialist position I am advocating is that, unlike the straightforward essentialism of Hegel and Marx, it does not invite us to regard succumbing to ideology as an all-or-nothing affair. Having given up on the straightforward essentialist goal of a revolutionary overcoming of ideology, it conceives of the opposition to ideology as an ongoing struggle.

The minimum we need from honesty, on this view, is that it should not be entirely powerless in the face of ideology. Even if it is possible to be sincere while in the grip of ideology, there will always be an opportunity to challenge such sincerity, on the basis of an appeal to honesty. This is not to claim that ideology can never be 'complete'. It seems perfectly possible for an ideology to attain the level of a comprehensive worldview, or system. The point is rather that however complete an ideology may be, it will always be vulnerable to honest critical appraisal – perhaps especially where it seems most complete, since it is at this point that it will most obviously fail to do justice to the complexity, and the conflicted nature, of the world of everyday experience.

The sceptical essentialist appeal to honesty will never compensate straightforward essentialists for what they are being asked to give up. The twentieth-century thinkers we have considered all give us, in their own ways, reasons to think that in their illusory systematicity every straightforward essentialism will be as fragile as a complete but unchallenged ideology. Nevertheless there is some reason to think that an uncompromising honesty will always be able to find a foothold, and enable us to unravel such systems from within, provided it is genuinely exercised.

I have characterised the failure involved in subscribing to ideological stereotypes as a failure of (collective) honesty. This understanding is clearly indebted to the Sartrean concept of bad faith, amongst other notions. But is it not still operating at the level of private perfectionism, rather than public morality as such? A commitment to honesty may involve essential reference to what would be accepted by an ideal community of enquirers, but wouldn't it still be, at bottom, a matter of realising one's own distinctive potential? Aren't the obligations founded upon it obligations to oneself, rather than to others?

The associated obligations *are* public obligations, and the Sartrean connection gives us the clue we need to see why this is the case. In some circumstances, falling short of the virtue of honesty involves not simply imperfection, but inconsistency. Consider again Sartre's account of sadomasochism in concrete relations with others. This syndrome is played out in the context of a 'game of mirrors'. In the conventional sadomasochistic pairing both the sadist and the masochist are aware of the dishonesty involved in the instrumental self-interpretation that the sadist deliberately induces in the masochist. It is not simply that neither party has good reason to believe that the masochist is a mere tool. (The view does not derive from mere ignorance.) Nor is it simply that both parties are aware that they do not have good reason to subscribe to the relevant instrumentalist view (it doesn't simply involve a 'leap of faith', in a situation characterised by a basic underdetermination – though again, if it were *only* that it would still represent a failure of honesty). Rather, the view involves positive mendacity, since both parties must simultaneously acknowledge that the masochist is no mere instrument. Sadomasochism of this kind then involves an extreme form of bad faith – not merely the bad faith that believes without good reason, but a form of bad faith that knowingly maintains incompatible commitments, for purposes of power and sadistic control. And this remains true whether or not the sadist's victims are themselves masochistic.

It is possible then to raise a sceptical essentialist ethic of respect for the *recognitional* capacities of persons from the level of an ethic of private perfection, founded on the fundamental virtue of honesty, to the level of an ethic of public obligation, which focuses in addition on the associated vice of mendacity. The theoretical burden accompanying such an ethic would reduce to this: that we grant that honesty as outlined above is a fundamental human virtue, which stands opposed to (but is not simply the logical complement of), a vice of implicit or explicit mendacity. If this is granted then we have the basis for a principled and public opposition to ideological stereotyping – either as the failure to exercise a suitably critical spirit with respect to the self-conceptions one adopts oneself, or induces others to adopt; or, when it takes Sartrean sadistic forms, as a positive mendacity with respect to interpreting

others. Indeed, we can imagine a Kantian-style maxim founded on these elements: *I should never endorse, and where possible should actively challenge, distorted self-conceptions – whether on my own part, or on the part of others.* This could be given as a formula for a form of respect for ontological persons that would go beyond respect for their interests and their capacity for self-determination. Adherence to this principle would involve taking seriously the idea of respect for the dignity of others, as the subject-objects of self-knowledge.

11.3 Implications for Professional Ethics

The foregoing discussion suggests that the idea of respecting human dignity, frequently invoked in applied and professional ethics – typically in connection with a principle of respect for persons, and more particularly a principle of respect for autonomy – requires a re-think.

As I suggested at the outset, the idea that the principle of respect for persons will serve as a foundational moral principle no longer looks tenable. Although we undoubtedly have duties to sentient beings who are not ontological persons, we cannot account for such duties on the basis of such a principle. It does not however follow that the principle of respect for persons no longer has any meaningful application. In addition to the aforementioned duties to sentient beings, we clearly have special duties to beings capable of some form of self-determination – though it may be that these duties differ, depending on whether the moral patient in question is of such a nature as to be merely capable of making choices, or is capable of making principled choices. (That is, whether it can meaningfully be characterised as autonomous, rather than simply 'self-determining'.)

But beyond this we also have special duties to ontological persons. One consequence of recent work in environmental and animal welfare ethics is to render the very idea that the class of moral patients and the class of moral agents might be co-extensive puzzling. Why, we might now reasonably wonder, would anyone ever have thought *that*? What is it about moral agents that would ever have seemed to justify the view that they are *uniquely* morally considerable? The idea that moral agents are uniquely morally considerable is, as I have suggested, no longer tenable. It does not however follow that moral agents have no special moral status. And in fact they do have a special moral status, though to appreciate this we have to understand the source of the peculiar dignity that attaches to them. This dignity does not reside simply in their capacity for moral agency, but in an associated capacity that is a frequently overlooked precondition of meaningful moral agency: the capacity for honest self-interpretation.

This capacity entitles those who possess it to special forms of respect and consideration. In addition to showing consideration for their well-being and their informed and considered choices, we have a duty to consider their capacity to be harmed and wronged, through being induced to adopt inadequate or distorting self-conceptions. To take this special vulnerability into account is to show respect for

the peculiar dignity of beings that are capable of understanding themselves, and by the same token capable of misunderstanding themselves. Such beings are capable of appreciating their own weaknesses and failings, as well as their own strengths; their dependency on others, as well as their relative self-sufficiency and autonomy. But they are also capable of misrepresenting themselves to themselves. And when they do this, their lives go correspondingly badly, irrespective of any additional consequences such misrepresentation may have.

Even those who are justifiably sceptical of the very idea of a human nature or essence must surely acknowledge that, insofar as we have a capacity for objective self-understanding, our flourishing is bound up with achieving such self-understanding. This drive to objective self-understanding is associated with a foundational virtue of honesty. It is also the basis of fundamental reciprocal responsibilities to others. If honesty with respect to my own self-understanding is fundamental to my own (sceptical essentialist) self-realisation, then I am bound to respect this capacity in others also, since I am essentially reliant on others for the ideal of objectivity that is a presupposition of the relevant form of honesty. My project of honesty with respect to myself is not a lonely and individualistic project then; it necessarily involves others. Possessing the capacity for self-understanding carries with it an imperfect duty to achieve an honest self-conception, and a perfect duty to refrain from mendaciously inducing distorting self-conceptions in others.

As generally binding as these responsibilities are though, they do not in practice fall on everyone in quite the same way. In our increasingly complex and technologically-driven world, expert knowledge is held in particularly high esteem. Such knowledge carries with it a great deal of authority, and this generates risks as well as benefits. Those in professional roles are often shown particular respect. But where those roles involve having expert knowledge of other human beings – whether in education, healthcare, law etc. – their incumbents are peculiarly exposed to the moral risks associated with inducing distorted self-conceptions. A few words from an acknowledged expert can consign a student/patient/client to a view of him/herself that echoes down subsequent decades, as an especially authoritative statement of who and what s/he really is. The consequences can of course be straightforwardly beneficial or harmful for the individual concerned – for his/her employment prospects, or mental health, for example. The self-conception that the professional induces in his/her client may be empowering and enabling, or it may be crushing and debilitating. But neither of these outcomes will be an unmixed benefit if that self-conception is at the same time a distorted one. Both a life marked by a mistakenly pessimistic self-conception, and a life spent in a fool's paradise, compromise the dignity of a being capable of better. The professional therefore has a particularly pressing responsibility to ensure that his/her knowledge is not deployed in a way that serves to induce or maintain distorted self-conceptions.

This means that, inter alia, the forms of self-understanding we induce in others must not be composed of ideological stereotypes, and must not be discriminatory, commodifying, or objectifying. There is perhaps a temptation to feel we are safe so long as we are merely truthful. But, as we have seen, this underestimates the extent of the challenge. A debasing self-conception may be composed of ideological

stereotypes, which are in turn composed of a good many truths. Telling nothing but the truth does not guarantee that we will not induce in others forms of self-understanding that put them in bad faith, and it may be, therefore, that honesty sometimes requires us to be cautious with the truth.

Particularly demanding versions of honesty (and associated concepts such as authenticity and bad faith) may be more suited to the practice of philosophy than to professional life. It is not inconceivable that as an intellectual or spiritual exercise we may wish to explore the possibility that all conceivable forms of self-understanding – all attempts to 'tell the truth about ourselves' – put us in bad faith. But at the same time it is possible to discern, in the context of our everyday personal and professional interactions with others, the broad outlines of what is and what is not acceptable. It is not only wrong to use another as a means, it is also wrong to treat them in a manner that leads them to understand themselves as a means, whether or not they are more likely to be treated as a means as a consequence. It is also wrong to interpret someone in terms of an ideological stereotype that clearly reflects and reinforces a regrettable history of oppression, or to induce them to understand themselves in such a way. Thus it may be wrong to advise someone, even with the best of intentions, to play on their gender or race and/or class background in order to secure a fair outcome in a given case. Equally, it may be wrong to encourage them to suppress a particular piece of information – for example regarding their sexual orientation. There will no doubt be contexts in which it is ultimately less dishonest for them to accept a label, and present themselves openly as (e.g.) gay. But there may be contexts in which thoroughgoing honesty requires us to encourage the same individual to reject any such label.

What is essential above all is to develop an awareness of the swarm of questionable self-conceptions that surround us and press in upon us, which only require an ill-advised but authoritative remark to become henceforth a structural component of an individual's self-understanding, to their permanent detriment. We have a duty not to make merely instrumental use of these self-conceptions, either in respect of ourselves, or of others. And this is not simply an affair of social psychology, or sociological theory: it is an affair of central moral importance, resting on well-grounded principles of reciprocal obligation. Although it is not the whole of ethics, this form of respect for persons belongs to the heart of ethics.

References

Aristotle. 1962. *Politics*. London: Penguin.

Foucault, Michel. 1970. *The order of things*. London: Tavistock.

Foucault, Michel. 1977. Nietzsche, genealogy, history. In *Language, counter-memory, practice*, ed. Donald F. Bouchard, 139–164. Ithaca, NY: Cornell University Press.

Heidegger, Martin. 1962. *Being and time* (trans: Macquarrie, J. and Robinson, E.). Oxford: Blackwell.

Kuhn, Thomas. 1962. *The structure of scientific revolutions*. Chicago, IL: The University of Chicago Press.

Marx, Karl. 1967. Economic and Philosophic Manuscripts (1844). In *Writings of the young Marx on philosophy and society*, eds. Loyd D. Easton and Kurt H. Guddat, 283–337. New York, NY: Anchor Doubleday.

Nietzsche, Friedrich. 1954a. On truth and lie in an extra-moral sense. In *The portable Nietzsche*, ed. Walter Kaufman, 42–47. New York, NY: Viking Penguin.

Nietzsche, Friedrich. 1954b. Thus Spake Zarathustra. In *The portable Nietzsche*, ed. Walter Kaufman, 112–439. New York, NY: Viking Penguin.

Nietzsche, Friedrich. 1966. *Beyond good and evil*. In *Basic writings of Nietzsche*, ed. Walter Kaufman, 179–435. New York, NY: Random House.

Nietzsche, Friedrich. 1974. *The gay science* (trans: Kaufman, W.). New York, NY: Vintage Books.

Popper, Karl. 1945. *The open society and its enemies*, Vol. 2. London: Routledge and Kegan Paul.

Rorty, Richard. 1989. *Contingency, irony and solidarity*. Cambridge: Cambridge University Press.

Rorty, Richard. 1991. Feminism and pragmatism. *Radical Philosophy* 59:3–14.

Wood, Allen W. 1990. *Hegel's ethical thought*. Cambridge: Cambridge University Press.

Index