

**CORPORATE PERSONALITY:
A POLITICAL THEORY OF ASSOCIATION**

PHD THESIS

Hans Kribbe
Government Department
London School of Economics and Political Science
Houghton Street,
London WC2A 2AE

UMI Number: U615608

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U615608

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

THESES

F

8165

1013051

ABSTRACT

This thesis aims to rescue the tradition of corporation theory from the implications of ontological and ethical individualism, which form important tenets of mainstream political theory. My argument for corporate personality builds on the Lockean and Humean theory of personal identity over time. According to this theory, personal identity is not rooted in the identity of a deep and indivisible Self, but in the relatedness between temporally distinct psychological stages. A person is a group of desires, beliefs, memories, traits and other attitudes, tied together by a causal string. On the best interpretation of the Humean view, personal identity has normative consequences. We are bundles but the string that ties the stages of one life together is formed by the normativity of action-guiding principles and by the mutual recognition of certain associative obligations. This normative claim is often meant to buttress ethical individualism against those who deny that persons exist at all. However, this thesis demonstrates that the claim also reinvigorates the idea of corporate personality. The argument develops along two different strands. First, it is shown that corporate personality is a political theory. It is the theory that compares co-operative relations between people with the co-operative relations between the stages of one person. For contractalist theories, the core virtue is justice. For corporate theories, the highest virtue is integrity. Second, corporation theory makes a real contribution to the field of political theory, in particular in an area where contractalism has traditionally encountered problems, to wit, the continuity of the contract. This thesis argues corporation theory is much more successful in explaining our transgenerational obligations to the past and future.

ACKNOWLEDGEMENTS

This is the fruit of what have been five years of relative isolation and not a few debts. I would like to express my gratitude to all those on whom I relied or who lent me a helping hand.

I should like to begin by thanking Paul Kelly, my supervisor, who pointed me in the right direction at the start and who helped me with sound advice on many occasions ever since. I also wish to acknowledge with thanks the financial support I received from the British Council and the Prince Bernhard Fund.

Many people took the trouble of reading and listening to my ideas at seminars, conferences and numerous other occasions. I have benefited enormously from their comments. I am particularly grateful to the participants of the Departmental Workshop in Political Theory at the London School of Economics, who have had to do so on several occasions.

I count myself fortunate in having had the excellent company of many good friends and colleagues. I should mention Daniel Rubenson, Jürgen De Wispelaere and Cillian McBride in particular.

One inevitably incurs many debts in life. But only few of them are impossible to repay. There are now three people to whom I owe such a debt. Two of them are my parents. The third is Beatriz.

INTRODUCTION. THREE JUSTIFICATIONS FOR CORPORATE PERSONALITY	6
CHAPTER I. THE STRANGE RESURRECTION OF A LONG FORGOTTEN ANALOGY.....	29
INTRODUCTION	29
1. THE PSYCHOLOGICAL PRINCIPLE OF PERSONAL IDENTITY	31
2. PERSON AND STATE	38
3. THE EXTREME CLAIM.....	43
3.1 <i>Prudence and Future Selves</i>	45
3.2 <i>Desert, Responsibility, Obligation and Past Selves</i>	48
3.3 <i>Justice and the Separateness of Persons</i>	49
CONCLUSION	53
CHAPTER II. INTEGRITY.....	55
INTRODUCTION	55
1. THE MODERATE CLAIM	58
2. THE PRACTICE THEORY	60
3. THREE OBJECTIONS TO THE PRACTICE THEORY.....	66
3.1 <i>Is Personal Identity Bifurcated?</i>	67
3.2 <i>Rule and Indirect Utilitarian Justifications for Person-based Ethics</i>	68
3.3 <i>Pluralism and the Practice Theory</i>	70
4. FROM IDENTITY TO INTEGRITY	72
4.1 <i>The Hierarchical Model</i>	73
4.2 <i>The Reason Model</i>	76
5. INTEGRITY AND MORALITY	81
6. INTEGRITY AND TIME.....	85
CONCLUSION	93
CHAPTER III. CORPORATE PERSONALITY: A POLITICAL THEORY	95
INTRODUCTION.....	95
1. THE REDUCTIONIST CONSTRAINT	98
2. SOVEREIGNTY CONSTRAINTS.....	108
2.1 <i>Conventionalism</i>	108
2.2 <i>Contractualism</i>	110
2.3 <i>Voluntarism and Obligations of Shared Intention</i>	116
CONCLUSION	124
CHAPTER IV. THE PERSON OF THE STATE	127
INTRODUCTION.....	127
1. PROBLEMS OF SCALE	128
2. COLLECTIVE INTENTION: AN INTERPRETIVE APPROACH	131
3. TWO UNPROMISING SOLUTIONS	135
3.1 <i>Liberal Nationalism</i>	137
3.2 <i>Liberal Impartiality</i>	141
4. INSTITUTIONAL PERSONS.....	142
4.1 <i>The Social Acceptance of Institutions</i>	146
4.2 <i>Why Integrity Still Matters at the Institutional Level</i>	149

4.3 <i>Pluralism and Boundaries</i>	155
CONCLUSION	159
CHAPTER V. EXPANDING OUR ETHICAL HORIZONS.....	162
INTRODUCTION	162
1. IRREDUCIBLY SOCIAL GOODS.....	164
2. PRUDENCE AND BEYOND: EXTENDING THE SCOPE OF REASON	166
3. CONSERVATISM	174
4. RECIPROCITY AND FUTURE GENERATIONS	179
4.1 <i>Generational Overlap</i>	181
4.2 <i>The Morality of Feeling</i>	183
5. WHAT, IF ANYTHING, IS WRONG WITH RESOURCE DEPLETION?.....	187
5.1 <i>The Veil of Ignorance</i>	193
5.2 <i>Kavka's Solution</i>	195
CONCLUSION	200
CHAPTER VI. WHY CARE FOR THE FUTURE?	202
INTRODUCTION	202
1. TRANSGENERATIONAL ASSOCIATION	203
2. TEMPORAL DISTRIBUTIONS WITHIN LIVES	206
3. TEMPORAL DISTRIBUTIONS WITHIN ASSOCIATIONS.....	215
3.1 <i>Society is Indeed a Contract</i>	217
3.2 <i>Why an Appeal to the Rights of Future People is Unnecessary</i>	221
CONCLUSION	227
CHAPTER VII. WHY CARE FOR THE PAST?	229
INTRODUCTION	229
1. WHETHER ONE GENERATION HAS A RIGHT TO BIND ANOTHER	232
2. IMPARTIALITY AND SPECIAL ALLEGIANCE	237
2.1 <i>One Thought Too Many</i>	239
2.2 <i>Justice, Expedience and Political Obligation</i>	242
3. ASSOCIATIVISM: ONE THOUGHT TOO FEW?.....	248
3.1 <i>The Conceptual Argument</i>	250
3.2 <i>Identity-dependent Arguments</i>	251
3.3 <i>Value-dependent Arguments</i>	252
4. BAD ASSOCIATIONS.....	256
CONCLUSION	266
CONCLUSION AND SUMMARY	268
BIBLIOGRAPHY.....	283

INTRODUCTION

THREE JUSTIFICATIONS FOR CORPORATE PERSONALITY

Most people believe that human beings have special ethical significance. Humans are moral persons, who have interests, rights and duties. This separates us from animals, mountains and oceans. However, most people do not believe that human associations are persons, with interests, rights and duties. The idea of a *corporate person* is commonly thought to be politically dangerous and metaphysically defunct. Individual and association, it is widely believed, are different moral entities. Individuals have basic standing in our ethics; associations do not. This thesis challenges that belief. It argues that corporate personality is alive. Indeed, it claims that corporation theory and ethical individualism spring from a common source, a single theory of the moral person.

Liberalism has become too individualistic. This was the central claim of the communitarian movement two decades ago. To restore the balance, communitarian theory claimed that language and culture and belonging were constitutive of our identities and of the human good. However, this claim does still not prove that associations have true ethical standing, as liberal theorists have since pointed out. Contemporary political theory, therefore, has learnt all it can learn from communitarianism and its rediscovery of culture and belonging. To appreciate the true ethical significance of associations, we need to look elsewhere. We need to turn to the tradition of corporation theory. This tradition agrees with ethical individualism in a basic respect. Persons are what ultimately matters. It suggests however that ethical individualism makes one, fatal mistake. It is wrong in thinking that only individuals may be persons.

I argue that corporation theory can be recovered without having to rely on implausible metaphysical assumptions. In fact, I claim that corporate personality and ethical individualism rely on the same ontological basis. To understand this, we first

need to focus on theories of personal identity. There is a school of thought that conceives of personal identity through time in terms of the relatedness between psychological states – thoughts, desires, beliefs, memories – at different moments. On this view, what we normally call a person is a mere *bundle* of experiences or psychological parts of a person. Opposed to the bundle theory stands another school of thought, which advances the position that a person is not a sequence of thoughts or psychological occurrences, but the subject of such mental experiences. On this view, a person is an *Ego*. That is to say, a person is that which owns certain thoughts and experiences, and therefore exists separately from them.

The latter theory of personal identity finds relatively little support among philosophers these days. Even so, it has been claimed that the Ego theory is implicit in our ethics. Hence, if it were false, this could have implications for our theories of practical reason. Indeed, Derek Parfit argues that the bundle theory undermines traditional accounts of prudence and rights-based moral theory.¹ According to Parfit, these accounts are viable only on the basis of the Ego theory, which he deems indefensible.

Others have challenged this position. From a Kantian viewpoint, they hold that we should begin from a conception of the person as a deliberative, accountable and temporally extended agent. Persons may be bundles of psychological states. But this view restores the temporal unity of agency, not by reviving the Ego theory, but by understanding this relatedness in terms of obligations and responsibilities of membership. To be related in the manner of a person, on this view, is to acknowledge past and future actions and interests *as your own*. Thus it is argued that the bundle theory, important though it may be in metaphysics, has no major consequences for practical reason.

¹ D. Parfit, *Reasons and Persons*, third and revised edition (Oxford: Clarendon Press, 1987).

Part of what this thesis argues is that these critics are right. We need to reject the idea that practical reason should become more impersonal. But more important than this is that in rejecting Parfit's argument, these critics get more than what they bargained for. The gist of what I want to demonstrate is that their arguments, while clearly individualist in spirit, facilitate a paradoxical move from ethical individualism to a form of ethical holism that includes the idea of corporate personality. To put it differently, while the bundle theory may not have the kind of consequences that Parfit and others envisage, it does significantly alter our perception of how the notion of a person fits into normative theory. And this change challenges certain important assumptions regarding the ethical standing of associations.

By way of introduction, I want to begin by saying something about the idea of a corporate person and what it entails. In this thesis, corporate personality refers to the idea that human associations qualify as persons in and of themselves, as so-called group persons, much in the same way that human beings qualify as individual persons.² To flesh out what this means, we should begin by specifying what the concept of a person refers to in general. Above all, it is important to emphasise the following two components.

First, the term person or personality is often used to indicate that a thing is a unit capable of having a good or a unified interest. By ascribing personality or personal identity to something we also denote that this thing holds rights and imposes legal and moral duties on others. In short, corporation theory puts forward a view on which human groups in general form the *loci* of interests and rights that are irreducible to, and may even outweigh the interests and rights of individual constituents.

² Throughout this thesis I use the term 'corporate' to refer to human groups, collectives or associations in general. The notion of corporate personality does not refer only, and not even principally, to the firm, but also to states, churches, trusts, foundations, nations, clubs, cultural groups, and so forth.

Second, the terms person and personhood are used to refer to agents and agency. When we describe something as a person, we point out, not just that it has unified interests, but also that it is capable of having and critically appraising a will, truly capable, therefore, of authorising action and speech. Corporate personality, in this more philosophical context, entails the view that associations are reflective intentional systems not reducible to the intentions of their individual members. Groups have a will of their own. Consequently, this group will may have greater authority than the member's individual wills.

Scruton sums up nicely what corporation theory implies:

All of the following can be said to be true of corporations (whether clubs, churches or firms):

- they make decisions;
- they act freely and responsibly;
- they have moral rights and duties
- they have legal rights and duties;
- they can make laws for themselves and their members, for the breach of which they are held responsible;
- they are objects of praise and blame; of loyalty, pride and affection; of anger, resentment and hate;
- they are historical beings, which flourish and decline according to the success of their undertakings;
- they have habits of mind, including moral virtues and vices;
- they stand in personal relations, and can adopt many of the roles adopted by human persons.³

In its many different manifestations, the idea that associations are somehow analogous to persons has, of course, a very rich and complex history in Western thought. Perhaps the most influential example is Plato's analogy in *The Republic* between virtue in the soul – reason ruling the appetites – and the just political order.⁴ The idea of a transindividual will also makes an important appearance in Rousseau, Hegel, Bradley and Bosanquet.⁵ It assumes an even greater role in Gierke's theory of

³ R. Scruton, 'Corporate Persons', *Proceedings of the Aristotelian Society*, 63 (1989), pp. 239-26, at p. 245.

⁴ Plato, *The Republic*, D. Lee, (ed.), second edition (London: Penguin Books, 1987).

⁵ J. J. Rousseau, *The Social Contract*, in: G. D. H Cole (ed.), *The Social Contract and Discourses* (London: Everyman, 1997); G. F. W. Hegel, *Philosophy of Right*, T. M. Knox, (ed.) (London: Oxford

the 'Genossenschaft' and the work of his British pluralist followers, Figgis, Cole and Laski.⁶

Given this long history, it is perhaps remarkable to see that corporation theory receives only little discussion in political theory today. The analogy between group and person has almost vanished, and if it appears at all, it is commonly greeted with suspicion or glee. Since the latter half of the twentieth century, at least, and in particular since the resurgence of contractual and rights-based approaches, political theorising has evidently been dominated by the principle of ethical individualism. This principle maintains that the notion of a person, either as a right-and-duty-bearing unit, an intentional agent or both, should only be applied to the individual. To attribute interests, rights and intentions to human groups, at any rate, is to lapse into philosophical error. Corporate personality, then, may boast a long tradition, but mainstream theory has resolutely consigned it to the dustbin of history. Thus, the idea belongs to an earlier époque and only merits examination in the context of the history of ideas. Theories of corporate personality have little to teach us today, as David Runciman confirms in his critical book *Pluralism and the Personality of the State*.

To be sure, this verdict may well be correct for some or even many of its manifestations; there are probably numerous particular theories of corporate personality that have little to recommend. Nevertheless, the idea that groups are persons is in my view too easily dismissed, by Runciman, and also by many others. It seems to me that corporation theory still contains a kernel of truth. What is more, I think it capable of making an important contribution to normative political theory. In any event, it is important to emphasise that in revisiting this notion, my thesis does not aim to provide another historical study of an ancient idea. Rather, its overall aim is to *revive* the notion

University Press, 1967); F. H. Bradley, 'My Station and Its Duties', in: *Ethical Studies*, second edition (Oxford, Clarendon Press, 1927), pp. 160-204; B. Bosanquet, *The Philosophical Theory of the State*, fourth edition (London: MacMillan, 1958).

⁶ For a discussion of the work of Gierke and the pluralists, see D. Runciman, *Pluralism and the Personality of the State* (Cambridge: Cambridge University Press, 1997).

of corporate personality, that is to say, to put it into a fresh guise, and to argue that, in other manifestations, it is still very much alive and even superior to ethical individualism. The ambition to rejuvenate and reconstruct corporation theory is also what sets this thesis apart from other work on the subject. By developing this constructive dimension my thesis adds a distinctive and relatively novel argument to political theory.

Consistent with this overall ambition, my approach is not historical but mainly analytical. I shall not be concerned to study what the 'dustbin' contains, how its contents have developed, and how they have ended up there in the first place. Interesting though this exploration may be, the priority is here to articulate an understanding of corporate personality that addresses the challenges set by contemporary political philosophy. Put differently, I want to investigate whether or not there is something in the idea of corporate personality that is still retrievable. And this means that the merits of this idea have to be assessed against the background of current debates in normative political theory and the individualist premise on which they are largely based.

However, to set out the challenge more clearly, it is useful to briefly look at the different arguments for corporation theory as they were advanced in the first part of the twentieth century. Received wisdom has it that there are only two justifications for corporate personality, the one juristic and the other philosophical. What is more, so it is usually argued, both justifications collapse under the weight of important objections and limitations. Essentially, my thesis argues for a third and irreducibly normative justification for corporate personality. However, I first want to discuss why the two standard accounts turned out to be dead ends.

The first justification for corporate personality focuses on the person as a legal right-and-duty-bearing unit. Legal persons are the subjects of rights and liabilities as defined

by the legal system.⁷ They may be represented in court, sue or be sued, enter in contractual engagements, own property, vote in elections, run for public office, claim welfare rights, and so forth. All mature legal systems recognise that individual human beings are legal persons in this sense. However, it is apparent that legal personality is usually not exclusively applied to individuals. More specifically, most legal systems also count human groups – companies, universities, cities and the state itself – as persons.⁸ Associations are summoned to court, they own property, and enter in contractual relations. Associations may not have the right to vote, they cannot marry, they do not have a right to health care, but at the same time it is evident that they hold certain other legal rights and obligations.⁹ We know, for example, that a university college could own property and enter into contractual transactions with other persons. No one seriously believes that the entitlements and obligations that are so acquired are ultimately divisible in separate parts belonging to the individual fellows. That would imply, absurdly, that every fellow owned part of the college's property and owed part of its debts on a personal title.

That said, while it is one thing to point out that associations are commonly recognised as legal persons in practice, it is quite something else to offer a coherent explanation for this fact. How could groups possibly qualify as persons? The *fiction theory* provides an answer to this question.¹⁰ It maintains that the legal system itself

⁷ P. A. French, *Responsibility Matters* (Kansas: University of Kansas Press, 1992), p. 134; S. J. Stoljar, *Groups and Entities: An Inquiry into Corporate Theory* (Canberra: Australia National University Press, 1973).

⁸ S. J. Stoljar, *Groups and Entities: An Inquiry into Corporate Theory* (Canberra: Australia National University Press, 1973). See also F. W. Maitland, 'Moral Personality and Legal Personality', in: H. A. L. Fisher (ed.), *The Collected Papers of Frederick William Maitland*, Volume III (Cambridge: Cambridge University Press, 1911), pp. 304-320.

⁹ For an account that traces the history of the idea of legal corporate personality back to Roman law, see P. W. Duff, *Personality in Private Roman Law* (Cambridge: Cambridge University Press, 1938).

¹⁰ For a discussion of the idea, see D. Runciman, *Pluralism and the Personality of the State*; M. Wolff, 'On the Nature of Legal Persons', *Law Quarterly Review*, Volume LIV (1938), pp. 494-521; J. Dewey, 'The Historic Background of Corporate Legal Personality', *Yale Law Journal*, Volume XXXV (April, 1926), pp. 655-673. For a more critical treatment: F. Hallis, *Corporate Personality: A Study in Jurisprudence* (Oxford: Oxford University Press, 1930), pp. 3-28; F. W. Maitland, 'Introduction', in: O. Gierke, *Political Theories of the Middle Age*, edited and translated by F. W. Maitland (Cambridge: Cambridge University Press, 1927), pp. vii-xlv.

determines who or what is to bear rights and obligations in its own name. And it is thought to be able to do so regardless of whether or not a being is a person in the philosophical and moral meaning of the term. Individuals are naturally perceived as legal persons since they are truly capable of having a will. But legal personality rests just as comfortably on the fiction of a will entertained by the legal system and its officials. According to this theory, in short, we need not look for whatever moral or psychological reality lurks behind the legal person. Legal and other meanings of the term 'person' refer to unrelated predicates. As Dewey once made the point: "What 'person' signifies in popular speech, or in psychology, or in philosophy or morals, would be as irrelevant, to employ an exaggerated simile, as it would be to argue that because a wine is called 'dry', it has the properties of dry solids; or that, because it does not have those properties, wine cannot possibly be 'dry'."¹¹

According to the fiction theory, legal personality should be compared with the term *persona* in the theatre. There it initially used to refer to the mask worn by an actor in a play. Later, the term was used simply to indicate the role or character of a particular part.¹² According to the fiction theory something similar could be said about legal personality, which is merely a mask or role in a juristic play, enacted by agents and created by an author. Legal persons, of course, must perform real actions and at some point we need real agency. However, groups need not be real agents themselves. They rely on individuals acting on their behalf. A group cannot literally sign a contract or transfer money itself. But the point is that these activities can be performed in its name. While its representatives act and decide, the consequences of their actions and decisions are assigned to the association.¹³ Conveniently, this construction allows us to contend

¹¹ J. Dewey, 'The Historic Background of Corporate Legal Personality', p. 656.

¹² P. W. Duff, *Personality in Private Roman Law*, pp. 3-6. However, that this metaphor ultimately proves only little is one of the problems in Runciman's *Pluralism and the Personality of the State*.

¹³ The fiction theory should not be confused with the so-called 'bracket theory' or 'aggregate theory'. This theory argues that while the legal rights of groups are in the end reducible to individual rights, corporate personality has a function as a technical device to diminish the complexity involved in such a reduction. The bracket theory is sometimes offered as a third justification of corporate personality, but it

two seemingly contradictory things. We can admit that only individual human beings are moral persons, but retain corporation theory. We could treat associations as right-and-duty-bearing units in the legal sphere, and yet let go of the view that they are real agents. Indeed, on this basis we could just as easily go on to award legal personality to animals, trees, ghosts and idols.¹⁴

However, the use of fiction has limitations in political theory. One problem is that we still have to explain how the state becomes a corporate person itself. For, as the author of fiction, the suggestion that the state is a legal fiction itself raises logical puzzles that are difficult to resolve.¹⁵ But a more fundamental problem is that while fiction creates *legal* personality for associations, it cannot produce *moral* rights and obligations. And corporation theory does not merely entail that groups figure as persons in positive law. It also includes the view that associations are valuable and worthy of respect. The idea suggests, in other words, that groups are ethical persons, specific units whose interests and liberties demand protection from a moral point of view.

The second kind of justification rejects the bifurcation between legal and moral concepts of the person. It suggests that both the individual and the group are legal persons precisely because they are real and moral persons. I refer to this kind of justification as *corporate realism* because, in contrast to the fiction theory, it calls attention to the reality that lurks behind the mask.

ultimately simply dismisses the idea. See S. J. Stoljar, *Groups and Entities: An Inquiry into Corporate Theory*, pp. 186-188.

¹⁴ S. J. Stoljar, *Groups and Entities: An Inquiry into Corporate Theory*, p. 2.

¹⁵ D. Runciman, 'What Kind of Person is Hobbes's State? A Reply to Skinner', *The Journal of Political Philosophy*, Volume 8 (2000), number 2, pp. 268-278, argues that on the Hobbesian theory the state is shown to be a fictional person, which relies on representation by the sovereign, rather than a real (or artificial) person relying on the representation by the sovereign. But the difficulty with this construction, apart from its ridiculous complexity, is that some real person will still have to authorise this fiction. The problem becomes visible when Runciman writes (p. 273) that the multitude 'create the conditions which allow the actions of the sovereign to be attributed to them as a single unit, since they are *jointly committed* [my emphasis] to taking responsibility for what the sovereign does.' But if 'the multitude' are capable of joint commitment, are they not thereby capable of joint action after all? See also Q. Skinner, 'Hobbes and the Purely Artificial Person of the State', *The Journal of Political Philosophy*, Volume 7, Number 1, 1999, pp. 1-29; D. Copp, 'Hobbes on Artificial Persons and Collective Action', *The Philosophical Review*, Volume 89, Issue 4 (October, 1980), pp. 579-606, especially pp. 603-606.

As an alternative to the fiction theory, corporate realism was perhaps most notably advocated by Otto von Gierke (1841-1921), mainly in the four volumes of *Das Deutsche Genossenschaftsrecht*, his magnum opus.¹⁶ Gierke's main objective was to defend the freedom and value of German associative life, which he saw originating in mediaeval guild society and the political theories of that era. Since then, the ancient notion of fellowship had been forced on a steady retreat by the natural law tradition. According to Gierke, the latter set of ideas encouraged rampant individualism on the one hand and an omnivorous sovereign state on the other, leaving only little room for civic life to flourish in the middle. The fiction theory, while attempting to retain corporate ideas in law, inevitably fell short in Gierke's view. It made corporate personality too vulnerable to the whims of the sovereign. Corporate realism wanted to restore the balance by pointing out that group personality existed independently of the law. As Barker put it: 'Legal group-personality is the shadow cast by real group-personality: it is the reflection of reality in the mirror of law.'¹⁷ The state may have been a corporate person too, but it had to recognise, in step with a federalist ideal, that it was a person made up of corporate persons.

Corporate realism faced a number of difficulties. Most importantly, since it relied on the notion of the group as an agent, it needed to explain how associations could have a will separately from the wills of their individual members, without relying on the aid of fiction or representation. Gierke sought refuge in a biological metaphor.¹⁸ A group mind, he thought, acts through the bodies of their individual members, just as the individual mind acts through the body's arms and legs. While not directly observable,

¹⁶ Only parts of this work, which was published between 1868 and 1913, have appeared in English. See O. von Gierke, *Political Theories of the Middle Age*, edited and translated by F. W. Maitland (Cambridge: Cambridge University Press, 1927); *Natural Law and the Theory of Society*, edited and translated by E. Baker (Cambridge: Cambridge University Press, 1934); *Community in Historical Perspective*, edited by A. Black, translated by M. Fisher (Cambridge: Cambridge University Press, 1990).

¹⁷ E. Barker, 'Introduction', in: O. Gierke, *Natural Law and the Theory of Society*, edited and translated by E. Baker (Cambridge: Cambridge University Press, 1934), p. lxvii.

¹⁸ See here O. Gierke, 'The Nature of Human Associations', edited and translated by J. D. Lewis, in: *The Genossenschaft-theory of Otto Von Gierke* (Madison: University of Wisconsin Studies in the Social Sciences and History, 1935), pp. 139-157.

the reality of a folk-soul, folk-feeling and folk-will could be deduced from the fact that there were communal laws and language, which he thought could not be explained without it. More importantly, according to Gierke, we could infer the reality of the 'corporeal-spiritual unity' of a Group Ego from its impact on our own consciousness.

The incorporation of our Ego in a social Being of a higher order is a matter of our own inner life. We are conscious of our self as a being enclosed in itself; but we are also conscious of our self as a part of a living whole which is operative in us. If we abstract our membership of our particular nation or State, our religious community and Church, our family and a variety of other groups and associations, we cannot recognise ourselves in the pitiable residue. But if we reflect on all these factors, we see that there is here no question of merely external bonds and fetters by which we are chained. It is a matter of psychical connections which extend down into our inmost being, and constitute integral parts of our spiritual existence. We feel that a part of the impulses which determine our activity proceeds from the communities by which we are permeated. We are conscious that we share in a life of community. If we derive from our internal experience a certainty of the reality of our Ego, this certainty is not limited to the fact of our being an individual life-unity; it also extends to the fact of our being a part-unity within the higher life-unities.¹⁹

In sum, on these realist foundations, corporate personality was closely linked with a robust version of metaphysical collectivism according to which groups had minds, thoughts and wills of their own.

We have here two very different justifications for corporate personality. The problem is that they both fall short for a variety of reasons. In point of fact, looking at corporation theory in law, it is apparent that the controversy between fiction theorists and realists has since long been dead. It is commonly thought that this debate ended in stalemate, that both theories were right in one sense and wrong in another, and that they made no difference to legal practice.²⁰ Hart drove the final nail in the coffin by arguing

¹⁹ O. Gierke, 'The Nature of Human Associations', cited in E. Barker, 'Introduction', in: O. Gierke, *Natural Law and the Theory of Society*, pp. lxxviii-lxxix.

²⁰ For example: S. J. Stoljar, *Groups and Entities: An Inquiry into Corporate Theory*, pp. 182-186; G. Teubner, 'Enterprise Corporatism: New Industrial Policy and the 'Essence' of the Legal Person', in: S. Wheeler (ed.) *A Reader on the Law of Business Enterprise* (Oxford: Oxford University Press, 1994), pp. 51-79. For further reconciliatory approaches: E. Barker, 'Introduction', in: O. Gierke, *Natural Law and the Theory of Society*; F. Hallis, *Corporate Personality: A Study in Jurisprudence*, pp. 239-246; A. W. Machen, 'Corporate Personality', *Harvard Law Review*, Volume XXIV, No. 4 (February, 1911), pp. 253-267.

that both theories, but realism in particular, were essentially mistaken about the nature of legal concepts.²¹ He famously suggested that legal concepts should be compared, not to fact or fiction, but to the rules of games.²²

Having said that, the debate on corporate personality has run a different course in political and ethical theory. There the idea of a group person means that associations form the *loci* of rights and duties, not just in legal practices, but also in moral and normative reasoning. Obviously, the fiction theory cannot deliver normative conclusions.²³ It is one thing to say that groups *can* be legal person, but it is quite something else to argue that they *should* have that status. The one is a claim about the legal order, the other a claim about the moral order. However, it is all too commonly surmised from this limitation that corporate personality could only be justified by realism. It is important to point out that this inference rests on a mistake. But *if* it were valid, and we still want to say something about the moral order, then we face a choice between two theoretical alternatives. We would have to either dismiss corporate personality altogether or claim instead that group persons are somehow metaphysically real.²⁴ Presented with this choice, it is not too difficult to see why corporation theory was completely dismissed.

The first alternative leads to the principle of *ethical individualism*. Briefly, this principle declares that only the individual enters into the Kingdom of Ends. To human associations it assigns the role of mere instrument. To be sure, ethical individualism does not rule out that associations are important, but the key is that their importance derives from the service they render to the individual's good.²⁵ The principle holds that

²¹ H. L. A. Hart, 'Definition and Theory in Jurisprudence', in: *Essays in Jurisprudence and Philosophy* (Oxford: Oxford University Press, 1983), pp. 21-48.

²² According to Hart, legal personality is an 'institutional fact'. For this last notion, see also J. R. Searle, *The Construction of Social Reality* (London: Penguin Books, 1996) and my discussion in Chapter IV. Notice also that, on this view, the same applies to the legal personality of individuals.

²³ See P. A. French, *Collective and Corporate Responsibility* (New York: Columbia University press, 1984), pp. 32-38.

²⁴ See F. W. Maitland, 'Moral Personality and Legal Personality', pp. 318-319.

²⁵ For example, M. Dan-Cohen, in his book *Rights, Persons, and Organizations: A Legal Theory for Bureaucratic Society* (Berkeley: University of California Press, 1986), argues that while organisations

the interests of the association are not in the final analysis normative or reason-giving. To put it differently, if somebody suggests that you have a reason – either of prudential or moral nature – to do *x*, then *x* must be justifiable in terms of your own well-being or that of other individuals. An appeal to the interest of the group as a whole – if this notion were intelligible at all – does not count as a reason for or against *x*. Whether measured in terms of preference satisfaction or pleasure or by some objective standard of goodness, only the good of an individual provides us with a reason for doing this or that. In Barry's words, the principle 'rules out appeals on behalf of God, Nature, History, Culture, the Glorious Dead, the Spirit of the Nation, or any other entity – unless that claim can be somehow reduced to terms in which only individual human interests appear.'²⁶

The principle of ethical individualism receives perhaps its most prominent expression in the rights-based theories of Rawls, Nozick, Buchanan and Dworkin.²⁷ Albeit with different substantive conclusions, these theories are united in their rejection of the classical utilitarianism of Bentham and Sidgwick, as has often been noted before.²⁸ Crucially, all point out that utilitarianism falls short precisely because it fails to acknowledge the full importance of the individual. While agreeing that the well-being of each counts *equally* in the calculus, utilitarianism fails to also attach significance to the *distinctness* of persons. For the utilitarian, it does not matter – not in a fundamental sense – whether utility is experienced by one individual as opposed to another. Since his guiding principle is maximisation, what matters more than anything is that utility,

could have rights, they could only derive them from consequentialist considerations. McMahan argues: 'organizations and other social entities – viewed as distinct from their members – are not appropriately accorded moral consideration in their own right.' See C. McMahan, *Authority and Democracy: A General Theory of Government and Management* (Princeton: Princeton University press, 1994), p. 62.

²⁶ B. Barry, 'Self-Government Revisited', in: *Democracy and Power* (Oxford: Clarendon Press, 1991), pp. 156-186, at p. 159.

²⁷ J. Rawls, *A Theory of Justice* (Oxford: Oxford University Press, 1992); R. Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974); J. M. Buchanan, *The Limits of Liberty: Between Anarchy and Leviathan* (Chicago: University of Chicago Press, 1975); R. Dworkin, *Taking Rights Seriously*, seventh impression (London: Duckworth, 1994).

²⁸ See for example, H. L. A. Hart, 'Between Utility and Rights', in: *Essays in Jurisprudence and Philosophy*, pp. 198-222.

pleasure or want-satisfaction is experienced. Thus, the problem is that utilitarianism places no checks on trade-offs between the utility of different persons. As Rawls points out, in neglecting such checks, it makes the major mistake of conflating the principle of rational prudence for one man with the principle of justice for society.²⁹ Indeed, utilitarianism conceives of society or of humanity in its entirety as one organism aiming to maximise its own want-satisfaction over time. The desires or experiences of different individuals are mistakenly treated as if they belonged to one super-person, who decides to sacrifice a lesser to a greater personal good.

In contrast to this view, Rawls and Buchanan aim to establish principles of distributive justice on the basis of the criterion of *general agreement*. Justice, they claim, is not grounded in the criterion of welfare maximisation, but in the more basic idea that distinct individuals have distinct and vital interests in the distribution of welfare or resources. Nozick advances a Lockean theory of justice on which individuals enjoy *natural rights* that serve as side-constraints on the maximisation of social welfare. And Dworkin bases his entire theory of justice on the principle that governments and law should always treat individuals with *equal concern and respect*.³⁰ Interestingly, he fleshes out this principle by introducing a distinction between ‘personal’ and ‘external’ preferences.³¹ A personal preference is a ranking of outcomes defined over what I could do or obtain. An external preference is a ranking of outcomes defined over what somebody else could do or obtain. According to Dworkin, the trouble with preference utilitarianism is that it does not shield the individual from the external preferences of other people, and he points to this failing in particular in order to explain why that theory infringes on the basic principle of equal concern and respect. To remedy this problem, Dworkin suggests that individuals have anti-utilitarian, trumping rights. While this may initially seem a matter of sophistication rather than a full rejection of

²⁹ J. Rawls, *A Theory of Justice*, pp. 22-33. A similar point is made by T. Nagel, *The Possibility of Altruism* (Princeton: Princeton University Press, 1970), p. 134, 138.

³⁰ R. Dworkin, *Taking Rights Seriously*, pp. 180, 272-274.

³¹ *Ibid.*, pp. 234, 275-278.

utilitarianism, the emphasis placed on the pivotal distinction between personal and external preferences evidently forces Dworkin back into line with the more obviously anti-utilitarian theories offered by Rawls, Buchanan and Nozick.

In sum, without neglecting their considerable internal differences, these theories all subscribe to one common assumption. Each individual person forms a distinctive unit of well-being, interests and preferences, and it is this unity that determines what we have reason to do, either morally or prudentially, not some greater whole of society, not even humanity as such.

The second alternative looks unappealing. This road leads straight back to the pitfall of realism. We would have to claim that collectives have preferences, intentions and various other mind-states. By choosing this path, in other words, we come to depend on something like the Ego theory of the person applied to the group. Like Gierke, we need to postulate a collective mind, existing underneath and entirely separate from the physical and psychological reality of individual human beings, and capable of deciding how its individual members should act. While it may be impossible to falsify such a theory, we also know – precisely, perhaps, because it is beyond refutation – that we cannot expect to marshal any serious support for corporate personality.

It is here that corporate personality is confronted by the objection of ontological reductionism and perhaps to a lesser extent also by methodological individualism. Like moths to the flame, ethical individualists are naturally drawn to both forms of reductionism, but it is important to note that these objections actually form different strands of individualism, which can in principle be detached from ethical and political concerns.³² Ontological reductionism claims that institutional and group action is always the aggregate of or the causal relatedness between individual actions and their (intended or unintended) consequences. We should always realise that collective

³² See for example C. McMahon, *Authority and Democracy: A General Theory of Government and Management*, pp. 52-69. For the many forms that individualism can take, see S. Lukes, *Individualism* (Oxford: Basil Blackwell, 1985).

intentions are the product of individual intentions. We habitually personify groups in everyday language, but this personification is never more than a convenient manoeuvre to avoid the staggering complexity of detail that collective action often involves. For example, the simple statement ‘Arsenal are looking for goals’ does not refer to a separate mental entity, which intends to commit its men forward. Rather, it is simply a figure of speech. This is how we sum up the dynamics in which individual players, in response to each other, decide to abandon their defensive positions, move in slightly different patterns, and so switch to a more attacking strategy. As J. S. Mill once put it, ‘Men are not, when brought together in society, converted into another kind of substance!’³³ Popper was perhaps the most famous advocate of the closely related principle of methodological reductionism. In his words, the principle stipulates ‘that all social phenomena and especially the functioning of all social institutions, should always be understood as resulting from the decisions, actions, attitudes, etc., of human individuals, and that we should never be satisfied by explanations in terms of so-called ‘collectives’ (states, nations, races, etc.).’³⁴ And Hayek, having brutally exposed the flaws of constructivism, was no less outspoken about ‘the even more primitive attitude which *personifies* such entities as society by ascribing to them possession of a mind’.³⁵

While methodological individualism forms the subject of ongoing debate, it is important to note that even in the communitarian literature we find few traces of corporate realism.³⁶ What we learn from Sandel, for example, is that communities are constitutive of our identities, and that we cannot, therefore, completely abstract ourselves from our roots and traditions. Similarly, Taylor makes the point that we all need a communal background of meaning in order to be able to become choosers at

³³ Cited in M. Gilbert, *On Social Facts* (London: Routledge, 1989), p. 408.

³⁴ K. R. Popper, *The Open Society and its Enemies*, Volume 2 (London: Routledge, 1945), p. 98.

³⁵ F. A. Hayek, *Law, Legislation and Liberty*, Volume 1 (London: Routledge, 1982), p. 27.

³⁶ For a history of methodological individualism, see L. Udehn, *Methodological Individualism: Background, History and Meaning* (London: Routledge, 2001). For a critical discussion, D. H. Ruben, *The Metaphysics of the Social World* (London: Routledge and Kegan Paul, 1985).

all.³⁷ But what we do not get from these theories is the view that groups and communities are persons in and of themselves. Pettit, in his misleadingly named book *The Common Mind*, holds to the typical view that, while our ability to think depends ‘horizontally’ on our relatedness with other individuals, our agency is not ‘vertically’ overruled by the intentions of collectives.³⁸ He claims to be a ‘holist’ rather than an ‘atomist’, but at the same time he supports ‘individualism’ rather than ‘collectivism’. Thus, while social relatedness is deemed to be crucially important to human agency, it is still the individual rather than the collective that is the agent. Anthony Black writes: ‘To think of groups as having some inner core (analogous to that of the human personality) is to bring down on oneself all the problems of holistic explanation, and a justifiable charge of anthropomorphism.’³⁹ And from this he concludes that an associational right is merely the kind of right that individuals have ‘*in virtue of their personal relationships or the groups to which they belong*’. His view resembles Kymlicka’s, who rejects the idea that ‘communities have rights independent of, and perhaps conflicting with, the rights of individuals who compose them.’⁴⁰ So while communitarianism clearly awards greater value to groups and associations than atomistic strands of liberalism, at a deeper level it still appears to be committed to individualism. The key problem, therefore, is that communitarianism does not explain why groups should have ethical standing comparable to that of individuals. This is precisely why so much of the communitarian argument is in the end diluted into ‘liberal-communitarianism’.

³⁷ M. J. Sandel, *Liberalism and the Limits of Justice* (Cambridge: Cambridge University Press, 1982); C. Taylor, *Philosophical Arguments* (Cambridge MA: Harvard University Press, 1995).

³⁸ P. Pettit, *The Common Mind: An Essay on Psychology, Society and Politics* (Oxford: Oxford University Press, 1993).

³⁹ A. Black, *State, Community and Human Desire* (Hemel Hempstead: Harvester Wheatsheaf, 1988), p. 131.

⁴⁰ W. Kymlicka, ‘Individual and Community Rights’, in: J. Baker (ed.), *Group Rights* (Toronto: University of Toronto Press, 1994), pp. 17-33, at pp. 18-19. For an alternative account of group rights along individualist lines, see also P. Jones, ‘Group Rights and Group Oppression’, *The Journal of Political Philosophy*, Vol. 7, No 4. (1998), pp. 353-377.

Corporate realism, of course, did offer such an explanation by claiming that associations have their own personality. But it is clear that this justification for corporate personality has lost its credibility. Human associations, social facts and collective actions are decomposable, if not to their non-social parts, then to the social network that connects the parts. And this, crucially, would appear to stamp out corporation theory. McTaggart, for example, claimed:

Compared with the worship of the state zoolatry is rational and dignified. A bull or a crocodile may not have great intrinsic value, but it has some, for it is a conscious being. The state has none. It would be as reasonable to worship a sewage-pipe, which also possesses considerable value as a means.⁴¹

As Nozick once put it:

Individually, we each choose to undergo some pain or sacrifice for a greater benefit or to avoid a greater harm: we go to the dentist to avoid worse suffering later; we do some unpleasant work for its results; some persons diet to improve their health or looks; some save money to support themselves when they are older. In each case, some cost is borne for the sake of the greater overall good. Why not, *similarly*, hold that some persons have to bear some costs that benefit other persons more, for the sake of the overall social good? But there is no *social entity* with a good that undergoes some sacrifice for its own good. There are only individual people, different individual people, with their own individual lives.⁴²

Corporate theory has thus been forced to the margins. Scruton still advocates the idea.⁴³ He hopes to revive corporate personality as an alternative to social choice theories and liberal contractualism. From an entirely different standpoint, the idea emerges also in work by Graham. His analysis of group personality seems to derive mainly from an interest in analytical Marxism, especially how it could benefit from a conception of class as a collective agent.⁴⁴ French defends a realist view on which

⁴¹ Cited in: A. Quinton, 'Social Objects', *The Aristotelian Society* Volume LXXVI (1975/76), pp. 1-27, at p. 13.

⁴² R. Nozick, *Anarchy, State, and Utopia*, pp. 32-33.

⁴³ R. Scruton, 'Corporate Persons'.

⁴⁴ K. Graham, 'Morality, individuals and Collectives', in: J. D. G. Evans (ed.), *Moral Philosophy and Contemporary Problems* (Cambridge: Cambridge University Press, 1987), pp. 1-18; *Karl Marx, Our Contemporary* (New York: Harvester Wheatsheaf, 1992). Graham, in his recent book *Practical Reasoning in a Social World: How We Act Together* (Cambridge: Cambridge University Press, 2002), has significantly extended his analysis to other groups and theories. He confronts the doctrine of the

organisations qualify as agents in virtue of having a ‘corporate internal decision structure’.⁴⁵ He argues that business corporations are morally and legally responsible for their actions. But these approaches struggle to convince us that something like a corporate person is possible at all.

This thesis argues that corporate personality really is possible. It aims to do so, not by returning to corporate realism, but by carving out a third justificatory theory of corporate personality. The negation of *ethical* individualism is implicit in the very idea of a corporate person. However, this is not necessarily true for *ontological* individualism. I have no quarrel with the latter strand of individualism. I subscribe to the view that associations are relational bundles of individuals without leaving any ontological residue. (The knack is that individuals are similarly reducible.) What I object to, however, is the claim that this reductionist ontology is incompatible with corporate personality.

I aim to provide a theory of corporate personality that frames this idea in terms altogether different from corporate realism or the fiction theory. I want to look for this justification in the view that the concept of a person, and especially the concept of the identity and distinctness of persons, is ultimately itself part of normative inquiry. Whereas some approaches in political theory hold that all normative considerations should be grounded in an antecedent view of the person, I argue, as Rawls has done in

distinctness of persons, in various interpretations, with the fact that individual actions tend to be causally interconnected. This leads him to suggest that the distinctness of persons is much less sharp than believed by most ethical individualists. According to Graham, that means that groups may play a significant part in moral theory, both as ‘moral agents’ and as object of concern or ‘moral patients’. I have been unable, at this late stage, to incorporate in this thesis the extensive discussion that his work here deserves. Suffice it to say that while some of Graham’s findings run parallel to my own, there remain what I think are important differences, specifically in how we arrive at our conclusions. First, I have tried to develop the notion of a corporate person by showing how it compares to the Lockean and Humean conception of personal identity through time. Second, I have given the notion of a person in general, and the idea of corporate personality in particular, a more explicit grounding in a normative – some would say conservative – conception of transgenerational continuity.

⁴⁵ P. A. French, *Collective and Corporate Responsibility* (New York: Columbia University Press, 1984).

later work, that conceptions of the person are not prior to practical reason.⁴⁶ Theories about what we have reason to do are validated or invalidated, not in relation to a conception of the person, but in reflective equilibrium with our considered judgements and intuitions. For my purposes it is more important to emphasise that normative theories – albeit often implicitly – *presuppose* a particular conception of the person and personal identity. Hence, to argue for a certain type of reason – prudence is a good example – is also to argue for a particular kind of personal unity. On this view, then, personal identity appears as a normative category, an obligation or a virtue. To put it differently, the conception of a person is an outcome of normative theorising, not an ontological theory about what we are.

As we shall see, Rawls and others have turned to this ‘practice theory’ to augment their own, largely individualist theories. However, along the same lines, I want to entertain the view that there are irreducibly normative justifications, not for ethical individualism, but for corporate personality. This practical justification thus moves away from both the fiction theory as well as the problem posed by ontological reductionism. For while it concedes that ontological reductionism about groups may well be true, it rejects the view that this has ethical individualism as its implication. It rejects this implication because it argues that corporate personality, like the notion of the individual person itself, is implicit in particular views and conceptions about what we have reason to do. Corporate personality *is* a view about what we have reason to do, and such views, I argue, are ultimately rooted in reflective equilibrium. In short, corporate personality is ‘political not metaphysical’.⁴⁷ On this theory, the notion of a corporate person is interchangeable with a particular view of associative obligation and responsibility.

⁴⁶ For example, J. Rawls, ‘The Independence of Moral Theory’, in: S. Freeman (ed.), *John Rawls: Collected Essays* (Cambridge: Harvard University Press, 1999), pp. 286-302.

⁴⁷ This Rawlsian turn of phrase was suggested to me by Paul Kelly.

Let me finally explain how the argument is structured. Chapter I introduces the bundle theory of personal identity. This theory holds that the diachronic identity of persons consists in a kind of relatedness or bundling between temporally distinct psychological states. Essentially, it suggests that the analogy between groups and persons holds up after all, but *inversely*. However, as a consequence, it could be claimed that less importance should be awarded to the good of persons, and more importance to the temporal stages of a person.

Chapter II concentrates on the individualist's reply to this view. Ethical individualists argue that, although personal identity is a relation, it is a relation which is in part constituted by a certain type of obligation and reason. To become a unified person you have to assume responsibility for your past and you have to recognise that your future selves have a bearing on your present reasoning. I refer to this response as the 'moderate claim'. Two further steps are important here. First, while Parfit uses a particular theory of personal identity to appraise certain theories of rationality and morality, I use practical considerations to appraise conceptions of the identity of persons. Second, the latter approach is fleshed out in terms of integrity.

Chapter III turns to the idea of corporate personality. To support this notion, the claims of the previous chapter have to be replicated at the associative level. First, on the basis of the practice theory, we have to demonstrate that the relatedness between persons has the same ethical significance as the relatedness between person-stages on the moderate interpretation. Second, we need to explain how we could apply the virtue of integrity, not only at the *intrapersonal* level, but also at the *interpersonal* level. This chapter concentrates purely on that second task and introduces the notion of a 'shared intention'.

Chapter IV argues more specifically that corporate personality and integrity could be applied at the larger scale of the state. It is one thing to argue that relatively small and egalitarian groups could come to share an intention. But it is something else to claim

that the model could be applied to institutionalised groups like states. Some of the problems are that states are highly complex organisations, that their membership and boundaries are sharply demarcated, and that they commonly harbour a great variety of different worldviews and identities. To cope with these three challenges, I argue that collective intention should be understood interpretively. I also argue that the corporate personality of the state is largely institutional.

Chapter V shifts the focus of the thesis to substantive normative issues. I argue that the concept of corporate personality, like the concept of an individual person, should be understood as a theory about what we have reason to do. But if theories of the person are theories about reasons, then we also have to make sure that corporate personality makes a telling contribution to normative theory. That is, it must be capable of doing normative work that ethical individualism leaves undone. This chapter explores a number of hypotheses to that effect and settles on the view that corporate personality explains how we are morally connected with past and future generations.

Chapter VI expands this hypothesis by looking in greater detail at the problem of future generations. To understand why we take future interests into account at all, I argue, we have to appeal to corporation theory.

Chapter VII, finally, investigates the normative relatedness between present and past generations, which plays a central role in the notion of political and associative obligation. Implicit in that notion is the idea that we are presently bound and obligated by the commitments of previous generations. I discuss and reject a model that frames such obligations in terms of our higher-order duties of impartiality. I then argue that the commitments of a previous generation are comparable to the commitments of a previous self. What this involves is a view of associative obligation that is not derivative from the higher source of duty, but from integrity.

To conclude, let me briefly reiterate what makes this thesis into an original contribution to the field of political theory. First, it recovers a previously defunct

tradition. Corporation theory was abandoned in the Twentieth Century because it was thought to depend on spooky metaphysics, which had moreover been linked with fascism and communism. This thesis shows that corporation theory and liberal individualism build on the same, non-metaphysical assumptions. It shows that corporate personality is ultimately rooted, not in the heady obscurantism of German nineteenth century romanticism, but in the clear-headed, Anglo-Saxon empiricism of John Locke and David Hume. Second, this thesis further demonstrates that corporate personality has a practical contribution to make to political theory, in particular in an area where ethical individualism and liberal contractualism have traditionally found it hard to make headway, to wit, the transgenerational continuity of the social contract. That is to say, corporation theory allows us to make better sense of our ethical and political obligations to the past and the future.

CHAPTER I

THE STRANGE RESURRECTION OF A LONG FORGOTTEN ANALOGY

...the true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other... In this respect, I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes of its parts.¹

David Hume

Introduction

Corporate realism attempts to assign to the state and other associations the ontological status commonly ascribed to the individual. That is to say, associations are depicted as intelligent organisms with minds and intentional states of their own.² Few contemporary theorists believe that this analogy makes much sense. I do not disagree. In this realist form, the idea of a corporate person should be rejected. The view that I discuss in this chapter, however, puts the analogy between association and individual on its head. I want to begin the case for corporate personality right at the opposite end of the spectrum. Granted, associations are not individuals 'writ large'. However, it could be argued that individuals are associations 'writ small'.

The reversal of this analogy, as I shall present it here, treads a fairly thin line between two affiliated claims. The first claim is that individuals have the same *ontological* status as associations. It is based on the extrapolation of the standard form

¹ *A Treatise of Human Nature*, second edition by L. A. Selby-Bigge and P. H. Nidditch (Oxford: Oxford University Press, 1978), p. 261.

² For such an organic theory of association, see O. von Gierke, 'The Nature of Human Associations'.

of ontological reductionism to a form that applies *tout court*, that is, to associative and individual persons alike. The second argument claims that, in consequence, individuals need to be awarded the same *ethical* status as associations. Together, both arguments show that ontological reductionism forms a double-edged sword for ethical individualism. On the one hand, it provides the tool that exposes the weakness of corporate realism. Collectives, of course, are not free-floating consciousness. They should be broken down in individually conscious parts. But on the other hand, a more thoroughgoing ontological reductionism uncovers that individuals are in turn like miniature associations. In other words, there is the possibility that reductionism is pointed at its own conceptual heart, namely that of the individual person as the apex of our rational and moral universe.

This possibility has been most notably advanced by Parfit in his *Reasons and Persons*.³ In Part Three of that book, Parfit sets out to achieve two main ambitions. First, he seeks to develop a particular ontological conception of the nature of personal identity. Second, he aims to show how this particular conception of identity undermines all person-based conceptions of rationality and morality, also under attack in the other parts of the book, and which claim or imply that personal identity is something that fundamentally matters.

To avoid any confusion, I should pause to point out in advance that I do not support Parfit's views across the board. I rely on them only in part. To be more precise, his second conclusion, that we should liberate morality and rationality from the idea of personality as such, cannot be sustained. In Chapter II, I shall explain why. But the premise of this conclusion strikes me as correct. That is, I find little at fault with the

³ My discussion focussed mainly on *Reasons and Persons*. Different statements of essentially the same argument can be found in: D. Parfit, 'Personal Identity', *The Philosophical Review*, Vol. 80, Issue 1 (January, 1971), pp. 3-27; 'On 'The Importance of Self-Identity'', *The Journal of Philosophy*, Vol. 68, Issue 20 (October 1971), pp. 683-690; 'Personal Identity and Rationality', *Synthese*, Vol. 53, No.2 (November 1982), pp. 227-241; 'Later Selves and Moral Principles', in: A. Montefiore (ed.), *Philosophy and Personal Relations: An Anglo-French Study* (London: Routledge and Kegan Paul, 1973), pp. 137-169; 'The Unimportance of Identity', in: H. Harris (ed.), *Identity: Essays Based on Herbert Spencer Lectures Given in the University of Oxford* (Oxford: Clarendon press, 1995), pp. 13-45.

claim that the personal identity of individuals is reducible to a bundle or web of constitutive parts in a way that mirrors the plight of associations.

I propose to elaborate this latter claim first. In Section 1, I sketch the Lockean and Humean background to Parfit's reductionism *tout court*. Next, in Section 2, I discuss the conflict between this view and the Ego theory. Section 3 claims that the reductionist criterion of personal identity corresponds in important ways to that of associations. Ontologically, it is argued there, individuals and associations are on a level pegging. Finally, I explore Parfit's ideas – arguably his most daring – on the ethical implications of this ontology of persons in Section 4.

1. The Psychological Principle of Personal Identity

In order to set out the backdrop to the argument, I want to start by drawing a broad distinction between two schools of thought in the philosophy of personal identity. The first school is inspired by the ideas of Locke and Hume, while its modern members include Grice, Parfit, Perry and Quinton. It advances what is often described as the psychological criterion of personal identity. According to this criterion, personal identity over time consists in the relatedness between a series of mental states at different moments. It is also referred to as the bundle theory. The second school is mainly inspired by the ideas of Descartes. It maintains that personal identity is grounded, not in a bundle of psychological states, but in the identity of the underlying substance of an Ego.

It is in accordance with Locke's own viewpoint that we should start with an account of what the concept or idea of a person essentially *means*.⁴ To be a person, according to Locke, is to be something that can only be described as a particular mental faculty or state of being. Briefly, it is to be a thinking and self-conscious system or 'a thinking

⁴ J. Locke, *An Essay Concerning Human Understanding*, Book II, Chapter 27, edited by P. H. Nidditch (Oxford: Oxford University Press, 1975), pp. 328-348.

intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing, in different times and places; which it does only by that consciousness, which is inseparable from thinking, and as it seems to me essential to it'.⁵ Locke focused on mental states and faculties because in his view the notion of a person was essentially a moral one. According to Locke, the term person 'is a Forensick Term, appropriating Actions and their Merit; and so belongs only to intelligent Agents capable of a Law, and Happiness and Misery.'⁶

Consequently, the crux of the Lockean concept of the person is that it severs the psychological aspects of personality from the substance or stuff that does the thinking. On the one hand, there are thoughts, beliefs, memories, or as Hume put it, 'perceptions', while, on the other hand, there is the being of which we might say that it 'owns' these mental states. There are, of course, distinct views as to what the nature of this being or stuff might be. Descartes, in his *Meditations*, proposed that there exists purely spiritual substance.⁷ Other views hold that there is only material substance. Locke believed in spiritual substance. In any case, he did not seek to dismiss this idea. Parfit and other Lockeans seem more partial to the materialist position. But whatever our view of the nature of substance might be, the decisive point for all Lockeans is that the idea of a person is separable from the idea of thinking substance. Ultimately, the person is constituted purely by conscious thought, by reflection and self-awareness, and emphatically not by whatever stuff does the thinking.⁸

Locke, then, first distinguished between the 'Idea of a Thinking Substance', the 'Idea of a Person', and the 'Idea of a Man'. By person is meant a self-consciousness, intelligent, reasonable and accountable being. By substance we mean the stuff that

⁵ Ibid., p. 335.

⁶ Ibid., p. 346.

⁷ R. Descartes, *Meditations on First Philosophy*, translated and edited by J. Cottingham (Cambridge: Cambridge University Press, 1996).

⁸ Thus, this bifurcation of ideas allows us the benefit of metaphysical agnosticism. See A. Quinton, 'The Soul', in: J. Perry (ed.), *Personal Identity* (Berkeley: University of California Press, 1975), pp. 53-72, especially pp. 53-58.

ultimately renders us capable of consciousness and thought. And by man we mean a living organism. Consistent with this tripartite division of ideas, Locke devised three distinct criteria of identity, each 'suited to the idea'. These criteria are not meant to settle what sort of things substance, man, or person are, that is, what separates them from each other. Rather, they decide what separates one substance from another, one man from another man, and one person from the next. What is at stake here is the principle of individuation, 'which determines a Being of any sort to a particular time and place, incommunicable to two Beings of the same kind.'⁹ To be more specific, Locke wanted to address how substance, man and in particular persons retain their identity *through time*. To say that a thing is identical in this sense is to point to a specific unity relation between that thing at t and at t_1 . And Locke argued that substance, man and person each have a different unity relation.

(1) *Perfect identity*. The identity of substance is identity in what Butler called 'the strict philosophical sense of the word'.¹⁰ A singular substance at t is identical at t_1 if, and only if, it continues to exist as numerically the same, unchanged and discrete particle. This criterion presumes that there is some indivisible substance, either of material or immaterial nature. But applied to a number of monadic particles, it implies, according to Locke, that if one particle were removed or added, it were no longer the same body or mass. A mass or body remains identical, however, if the same parts were 'differently jumbled'.¹¹

(2) *Organic identity*. The idea of man refers to a living body of which the parts are organised so as to sustain a continued life. In this sense, the idea of a man is not different from the idea of animal or vegetable. Both are living and growing organisms. For Locke, the identity of a man is the continuation of the same life 'by constantly fleeting particles of matter, in succession vitally united in the same organized body'. So,

⁹ J. Locke, *Essay Concerning Human Understanding*, p. 330.

¹⁰ J. Butler, 'Of Personal Identity', in: J. Perry (ed.), *Personal Identity*, pp. 99-105, at pp. 100-101, originally in *The Analogy of Religion* (1736).

¹¹ J. Locke, *Essay Concerning Human Understanding*, p. 330.

while the parts can be replaced, a living body remains identical as long as the body parts sustain the same life.

(3) *Personal identity*. The identity of a person refers to the ‘sameness of a rational being’. But in what consists this sameness? According to Perry, we first need to distinguish between the two sub-questions to fully appreciate the complexity of the problem. Where K stands for person, we have to determine:

1. What relation obtains between simultaneous K-events that are events belonging to the same K;
2. What relation obtains between K-stages that are stages of the same K.¹²

The first question is less important. It probes into the *synchronic* nature of a person at any one moment in time. What makes me one person at this very moment? As we saw, for the Lockean, a person is constituted by a variety of psychological experiences or what Hume called perceptions. On one view, we could have just one such experience at a time, so the question does not arise. We are a thought, desire, perception, or some other ‘person-event’. On another view, such events first need to be bundled in a synchronic unity relation. But both views eventually arrive at the notion of a temporal part of a self or a ‘person-stage’.¹³ The second question examines the *diachronic* unity of a person over time. What historical relation should hold between two or more person-stages such that we could say that these person-stages form one person with one personal history? Is there a special connection such that if and when person-stages are so connected, certain psychological experiences at t and t_1 are the experiences of the same person?

According to Locke himself, the faculty of consciousness was the key to both questions. It is through consciousness, he argued, that simultaneous sensations, experiences and thoughts are identifiable as mine. If I am thirsty and tired these experiences are mine because I am aware of them. These events are suitably linked into

¹² J. Perry, ‘The Problem of Personal Identity’, in: J. Perry (ed.), *Personal Identity*, pp. 3-30, at p. 9.

¹³ For a comparable notion, see H. P. Grice, ‘Personal Identity’, in: J. Perry (ed.), *Personal Identity*, pp. 74-95; at p. 86, ff. See also the notion of a ‘soul-phases’ in A. Quinton, ‘The Soul’, p. 59.

the same person-stage, that is, because they figure in the same state of consciousness. More importantly, Locke went on to claim that personal identity is extended through time by means of memory.

This personality extends it *self* beyond present Existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it *self* past Actions, just upon the same ground, and for the same reason, that it does the present.¹⁴

If I remember visiting New York in 1996, then I am consciously aware that I then visited New York. In other words, I know this experience to have been mine.¹⁵ Thus, my person-stages are linked into the same personal history if a later person-stage contains or could (if I made a proper effort) contain a memory of an experience of an earlier person-stage.

This memory criterion invited much early critique, most notably by Joseph Butler and Thomas Reid.¹⁶ In response to this, others have extended and fine-tuned Locke's argument in diverse and elaborate ways. While it goes beyond the scope of this thesis to address all these developments, it is important to mention the following two. First, as Hume points out, personal identity should include connectedness between all kinds of psychological states.¹⁷ Personal identity does not just consist of memory, but also of thoughts, beliefs, intentions, desires, character traits, and so on. This extension is important because the idea that memory is a necessary rather than a sufficient criterion for personal identity is in itself implausible. I do not now remember many of my experiences last year. Consequently, the basis of my personal identity would seem indefensibly weak. On the extended version of the criterion, however, my identity is

¹⁴ J. Locke, *Essay Concerning Human Understanding*, p. 346.

¹⁵ *Ibid.*, p. 336.

¹⁶ J. Butler, 'Of Personal Identity'; T. Reid, 'Of Identity', pp. 107-112; 'Of Mr. Locke's Account of our Personal Identity', pp. 113-118, both in J. Perry (ed.), *Personal Identity*, originally in *Essays on the Intellectual Powers of Man* (1795). For a discussion of the idea of personal identity in a historical context, see R. Martin and J. Barresi, *Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century* (London: Routledge, 1999).

¹⁷ Hume discusses the issue of personal identity in *A Treatise of Human Nature*, Book I, Part IV, Section 6, pp. 251-263. For this extension of the criterion, see especially pp. 260-262, and also D. Parfit, *Reasons and Persons*, p. 205, 208.

more robust. I may not remember much about last year, but I have the same character traits and mostly the same beliefs and long term plans. Hume argues that two person-stages are linked into one personal history whenever they sufficiently *resemble* each other. So if a desire or some other mental characteristic at t is still present at t_1 , then both person-stages are connected in the relevant way. He also suggests that person-stages are relevantly linked by *causal* connections. As Shoemaker puts it, two person-stages are linked 'if the later of them contains a psychological state (a memory impression, personality trait, etc.) which stands in the appropriate relation of causal dependence to a state contained in the earlier one.'¹⁸ So even if a desire or a conviction is not retained in a later person-stage, it may still have inspired or formed the source of a later desire or conviction. On this view, then, person-stages are linked in the way of a person, not just if there are connected by memory, but also if they resemble each other or if a person-stage contains a 'successor state' of an earlier self.

Second, building on the work of others, Parfit makes a distinction between psychological connectedness and psychological continuity. Connectedness refers to the direct psychological links between two person-stages. If I could remember an experience I had twenty years ago, then there is one such link. Locke seems to have thought that connectedness is a necessary condition of personal identity through time. However, as his critics noted, psychological connectedness is not a transitive relation.¹⁹ Neo-Lockeans have since tried to amend this criterion in order to overcome this difficulty. The solution usually advanced grounds the criterion in the continuity of mental states rather than just the direct connectedness between them.²⁰ Parfit points out that continuity refers to an overlapping chain of strong psychological connections. On

¹⁸ S. Shoemaker, 'A Materialist Account', in: S. Shoemaker and R. Swinburne (eds.), *Personal Identity* (Oxford: Basil Blackwell, 1984), pp. 67-132, at p. 90, and the subsequent discussion of functionalism at pp. 92-97.

¹⁹ See T. Reid, 'Of Mr. Locke's Account of our Personal Identity', pp. 114-115, and the discussion in D. Parfit, *Reasons and Persons*, p. 206.

²⁰ See H. P. Grice, 'Personal Identity'. A. Quinton, 'The Soul', sets out a similar idea explaining the relatedness between 'soul-phases'.

this view, each connection is like a strand in a rope. While most strands do not touch each other directly, they are all indirectly connected into one rope. Hence, Parfit defines the criterion of personal identity – or what he refers to as ‘Relation R’ – in terms of ‘psychological connectedness and/or psychological continuity’.²¹

Briefly, I now want to turn to the second school of thought. At the deepest level, this school rejects the Lockean distinction between the identity-criteria for person and substance. The identity of a person, in fact, is not given by the temporal continuity of purely mental states such as thoughts, intentions and memory, but rather by the persistence of an underlying substratum, an Ego, to which these thoughts and intentions are said to belong. Personal identity through time is not the identity of thought and perception, but the identity of the subject that thinks and perceives. As Reid expressed the difference between Locke’s view and his own:

My personal identity [...] implies the continued existence of that invisible thing which I call *myself*. Whatever this self may be, it is something which thinks, and deliberates, and resolves, and acts, and suffers. I am not thought, I am not action, I am not feeling; I am something that thinks, and acts and suffers. My thoughts, and actions, and feelings change every moment; they have no continued, but a successive, existence; but that *self*, or *I*, to which they belong, is permanent, and has the same relation to all the succeeding thoughts, actions and feelings which I call mine.²²

In other words, thoughts must belong to a Thinker, something that exists separately from our mind-states.²³ The most important upshot of this theory is that personal identity is irreducible. Persons are not bundles of the temporal parts of a person, tied together as the strings in a rope, but indivisible. As Reid put it: ‘The identity of a person

²¹ D. Parfit, *Reasons and Persons*, p. 216. A further objection holds that the Lockean criterion *presupposes* personal identity. For example, if I remember a past experience, then we presume that I had that past experience. But this presumption may be false. For solutions, see for example D. Parfit, *Reasons and Persons*, pp. 220-222; S. Shoemaker ‘Persons and their Past’, *American Philosophical Quarterly* 7 (1970), pp 269-285; J. Perry, ‘Personal Identity, Memory, and the Problem of Circularity’, in: J. Perry (ed.), *Personal Identity*, pp. 135-155; S. Shoemaker and R. Swinburne (eds.), *Personal Identity* (Oxford: Basil Blackwell, 1984), pp. 80-88.

²² T. Reid, ‘Of Identity’, p. 109.

²³ For a contemporary defence, see R. Swinburne, ‘Personal Identity: the Dualist Theory’, in: S. Shoemaker and R. Swinburne (eds.), *Personal Identity*, pp. 1-66.

is a perfect identity: wherever it is real, it admits of no degrees; and it is impossible that a person should be in part the same, and in part different; because a person is a *monad*, and is not divisible into parts.²⁴

2. Person and State

I do not want to concern myself here with the debate between both theories. I shall simply assume that the reductionist account along psychological lines is the more plausible theory.²⁵ I now want to flesh out the psychological approach by looking at some of its most remarkable upshots. That is, if the bundle theory were correct, individuals would after all be comparable to states and other associations. I already remarked that this (inversed) analogy has two different aspects. First, it claims that the identity-relations of persons are like the identity-relations of associations. Second, it argues that we should consequently treat persons like we normally treat associations in normative theory. I want to concentrate on the issue of identity first.

Reductionism about personal identity turns the analogy between person and state upside down. Like Hume, Parfit argues:

²⁴ T. Reid, 'Of Identity', p. 111.

²⁵ Brian Garrett, in his book *Personal Identity and Consciousness* (London: Routledge, 1998), distinguishes between no fewer than seven forms of reductionism. Here, I make do with a basic form, shared by all models, according to which 'we should see the ontological status of persons as somehow secondary or derivative relative to the supposedly primary elements of our ontology' (p. 25). I should further point out that there is a third school of thought, which is also reductionist, but places a greater emphasis on the physical nature of a person's existence. See for example J. McDowell, 'Reductionism and the First Person', in: J. Dancy (ed.), *Reading Parfit* (Oxford: Blackwell Publishers, 1997), pp. 230-250, who accuses Parfit of advancing an implicit form of Cartesianism. Briefly, this school claims that psychological states are in the end physical states of our bodies or specific parts thereof, and, inasmuch this is true, the continuity of these physical states is a necessary part of the criterion of personal identity. Philosophers often use thought experiments like brain transplantation and tele-transportation to tease out our intuitions in these matters. Clearly, it goes beyond the scope of the thesis to look at this controversy at length. What is more, it has been suggested that our intuitions in this matter are essentially unclear. See B. Williams, 'The Self and the Future', in: J. Perry (ed.), *Personal Identity*, pp 179-198. I agree with Rovane, who, in her book *The Bounds of Agency: An Essay in Revisionary Metaphysics* (Princeton: Princeton University Press, 1998), argues that if we start from the ethical point of view of persons as intelligent and therefore deserving and accountable agents, we should incline towards the psychological criterion. While mental capacities may depend on the body and brain, the latter matter only contingently from this ethical standpoint.

When considering nations, most of us are Reductionists. We believe that the existence of a nation involves nothing more than the existence of its citizens, living together on its territory, and acting together in certain ways. In contrast, when considering persons, most of us believe the Non-Reductionist view. We believe that our identity must be determinate. This cannot be true unless a person is a separately existing entity, distinct from his brain and body, and his experiences. Most of us are thus Reductionists about nations but not about people. It is the difference between these common views which explains the two comparisons. The claim that X is like Y typically assumes the common view of Y. We shall therefore say, 'People are like nations' if we are Reductionists about both. If we are Non-Reductionists about both, we shall say, 'Nations are like people'. The belief in super-organisms may be a Non-Reductionist View about nations.²⁶

The nub is that psychological reductionism cannot account for personal identity in the strict or perfect sense of the term. Personal identity thus becomes comparable to the identity of political parties, churches, nations, states and other relational facts. Like the identity of associations, the idea of personal identity must be applied in the *imperfect* sense.

As we saw, perfect identity is identity of the kind that admits neither of degrees nor change. Whether a thing is perfectly identical at another time is a question that has a determinate answer. A_t is identical to $A_{t'}$ in the perfect sense if, and only if, they are *numerically* the same thing. On the Ego theory, personal identity is perfect, a matter of all or nothing, not of degree. However, if the person is constituted by psychological events, this is no longer true. Reid wrote: 'Our consciousness, our memory, and every operation of the mind, are still flowing like the water of a river, or like time itself.'²⁷ Thus, psychological events form a stream of incessant change and movement. My thoughts, feelings and ideas now are never numerically identical to my thoughts at the next moment. They might be matching in content, or they might be connected in a causal chain, but they are not literally the same thoughts. So once the bifurcation

²⁶ D. Parfit, *Reasons and Persons*, p. 332. For a similar analogy, though not grounded in this conception of personal identity, see G. S. Kavka, 'Is Individual Choice Less Problematic than Collective Choice?', *Economics and Philosophy*, 7 (1991), pp. 143-165.

²⁷ T. Reid, 'Of Mr. Locke's Account of our Personal Identity', p. 116.

between the ideas of substance and a person goes through, personal identity, in this perfect sense, must be a fiction.

We have to turn, therefore, to a relaxed understanding of identity, a notion more forgiving of change and degrees. As Hume noted, many of our everyday judgements on the identity of objects are relaxed in this sense. For example, when I shave off my beard, I do not conclude that my body is no longer the same. We believe that such a partial change is simply too insignificant to warrant this conclusion. As the bulk remains intact, so my body endures. Hume mentions the example of a river. The water molecules that form that river are forever in flux, hurtling down to the sea, eventually merging with it within days or weeks. Yet we do not hesitate to refer to it as the same river. Even rapid motion may not affect our sense of the identity of things. What these examples show is that our sense of identity survives in the face of change as long as this change is 'identity-preserving'.²⁸ And this means that changes have to respect certain constraints. It matters, for instance, that change is gradual or insignificant in proportion to the whole. Usually, it also matters that the replacing parts perform the same function or are causally related to the previous parts. Such constraints form the basic elements of a theory of imperfect identity.

Essentially, Hume and Parfit apply this understanding of imperfect identity to persons.²⁹ Relation R refers to the extent to which some person-stage Jack at t is psychologically connected and/or continuous with some other person-stage John at t_1 . But this entails that the identities of persons are now comparable to those of political parties, states and other relational facts. For example, we know that the identity of

²⁸ I borrow the terminology from D. H. Ruben, *The Metaphysics of the Social World*, pp. 63-38.

²⁹ Having said that, the conclusion that Hume ultimately reached, in the Appendix to *A Treatise of Human Nature* (pp. 633-636), was sceptical. The entire notion of personal identity was in the end merely a fiction. The fact that people assigned identity to changing things merely meant that we were prepared to deceive ourselves. We merely conjectured identity by 'a customary association of ideas'. Hume agreed with Locke that mental states or 'perceptions' were 'distinct existences' that could 'form a whole only by being connected together'. However, for him, it was unclear how such connections could ever be discovered or perceived, as he also generally held that '*the mind never perceives any real connexion among distinct existences.*'

groups and associations is *indeterminate*. Is France the same country as it was fifty years ago? While some things have changed, other things have not. Is this change identity preserving? Here the problem is that there is no non-arbitrary cut-off point. Whether France is the same cannot be decided with accuracy. We know that there is some measure of historical continuity. But to say that France is or is not the same country is arbitrary. The same is now true for personal identity. According to Parfit, the question whether or not some person is identical through time has no determinate answer.³⁰ The psychological connectedness between person-stages tends to vary. Remote person-stages are often less connected. Indeed, they may no longer be connected at all. As he puts it:

On my proposed way of talking, we use 'I', and the other pronouns, to refer only to the parts of our lives to which, when speaking, we have the strongest psychological connections. When the connections have been markedly reduced – when there has been a significant change of character, or style of life, or of beliefs and ideals – we might say, 'It was not *I* who did that, but an earlier self'.³¹

Do two such person-stages still belong to the same person? Contrary to what people tend to believe, the answer is not simply 'Yes' or 'No'. The only answer available is that they are to some degree psychologically connected. We could say that they are either weakly or strongly connected, but we cannot conclusively decide that they are identical or not.

We also know that associations often split up. Belgium may divide in a Flemish part and a Walloon part. In which part, if at all, will Belgium continue? Or does it continue in both? These seem hard questions. But they are equally hard for personal identity. One consequence of the psychological criterion of personal identity is that persons could branch or divide in two.³² In such cases, a person-stage Jack at t has strong psychological connections with John at t_1 and with Jim at t_1 . Personal division may be

³⁰ D. Parfit, *Reasons and Persons* pp. 229-243

³¹ *Ibid.*, pp. 304-305.

³² *Ibid.*, pp. 245-266.

technically impossible, but it is not inconceivable. Parfit imagines a world where surgeons could separate the two halves of Jack's brain, each half retaining consciousness. Both halves could then be transferred to different bodies resulting in the birth of John and Jim. Both John and Jim are psychologically connected with Jack in the way of a person. They have memories of Jack previous experiences, they share his beliefs and values, and so on. Now, on the non-reductionist theory, persons cannot divide into parts. However, on the basis of the reductionist theory, Jack can be seen to survive as both John and Jim.³³ According to Parfit, personal division is 'almost as good as survival' because R-relatedness is what fundamentally matters. So he claims that 'Since my relation to each of the resulting people is about as good as if it were identity, it may carry most of the ordinary implications of identity.'³⁴ For example, John and Jim may both be held responsible for Jack's actions. The key is that Jack and John and Jim are all psychologically connected. There is simply nothing we could sensibly add.

On the reductionist account, in short, the identity of a person is pulled together from temporally separate parts. Hence, it should be obvious that a person cannot be a monadic unity. The unity of a person is a measure of organisational integrity. Comparing persons to associations, Parfit writes:

In 1881 the French Socialist Party split. What happened? Did the French Party cease to exist, or did it continue to exist as one or other of the two new Parties? Given certain details, this would be an empty question. Even if we have no answer to this question, we could know just what happened.³⁵

And elsewhere he argues:

Talk about successive selves can easily be misunderstood, or taken too literally. It should be compared with the way in which we subdivide a nation's history. We call this the history of successive nations, such as Anglo-Saxon, Medieval, and Tudor England.³⁶

³³ Ibid., pp. 258-259.

³⁴ Ibid., p. 271.

³⁵ Ibid., p. 260.

³⁶ Ibid., p. 306.

Persons, in other words, should not be contrasted with groups. We are merely associations of a particular kind.

3. The Extreme Claim

Parfit's analysis puts associations and individuals back on a par. Associations are constituted by numbers of individuals bound together by a particular relation. And this means that associations have no identity in the strict sense. However, this fact does not differentiate associations from persons. For, as we have seen, persons are in turn constituted by a number of person-events and person-stages bound together like the strings in a rope.

Now, of course this is not the kind of analogy that corporate realism had hoped to uncover. Rather than raising the moral stature of associations, the analogy may only show that we ought to demote the moral standing of persons across the board. It may prove, in other words, that individuals ought to be treated like we normally treat associations, rather than the other way around. To put it differently still, this analogy could now be used to argue that ethical individualism is just as old-fashioned as corporate personality.

Locke's Eighteenth Century critics thought that this would inevitably be the outcome. Reductionism about personal identity, that is, would inevitably result in moral bankruptcy. They therefore thought that Locke had to be wrong. 'A part of a person', Reid wryly remarked, 'is a manifest absurdity.'³⁷ Even the briefest of introspective moments or the tiniest morsel of common sense would surely indicate that much. For Reid, strict personal identity was the 'foundation of all rights and obligations'.³⁸ For if Locke were right, he thought, 'no man could be responsible for his actions.'³⁹ Butler

³⁷ T. Reid, 'Of Identity', p. 109.

³⁸ *Ibid.*, p. 112.

³⁹ T. Reid, 'Of Mr. Locke's Account of our Personal Identity', p. 117.

pointed out that Locke's theory was equally hostile to prudential reasoning: 'for if the self or person of today, and that of tomorrow, are not the same, but only like persons, the person of today is really no more interested in what will befall the person tomorrow, than in what will befall any other person.'⁴⁰ Put differently, Locke threatened to explode all the familiar reasons that guide human action. Concepts like desert and punishment would, sure enough, lose their last shred of authority. The choice we faced, according to this objection, was between reason on the one hand and moral anarchy on the other.

Now, whatever he meant to say, Locke did certainly not want to leave morality in ruins. However, note that for the hardnosed reductionist these practical objections merely beg the question. According to Parfit, indeed, the objection only shows how much practical reason is tainted by Cartesianism.⁴¹ Parfit claims that we should simply revise our theories of rationality and morality. Having said that, the practical upshots of reductionism are by no means self-explanatory. Parfit makes an important distinction between the 'extreme claim' and the 'moderate claim'. Both claims are essentially at odds over how to appraise the moral and rational status of R-relatedness. The extreme claim bluntly denies that relational facts could have moral and rational significance. Thus, it denies that the interests and intentions both of past and future person-stages are in any sense relevant to my present self. And by doing so it denies exactly what the moderate claim wants to affirm.

According to Parfit, either claim is in principle defensible and thus his own positioning remains obscure, straddling both views.⁴² When confronting the effects of reductionism on prudence, responsibility, punishment and obligations of commitment he adopts a decidedly moderate stance. However, there is less trace of this moderate stance in his treatment of justice. While it is clear that Parfit's overall agenda is to show

⁴⁰ J. Butler, 'Of Personal Identity', p. 102.

⁴¹ D. Parfit, *Reasons and Persons*, p. 324.

⁴² He writes on p. 312 of *Reasons and Persons*: 'Though we are not forced to accept the Extreme Claim, we may be unable to show that it should be rejected. There is a great difference between the Extreme and Moderate claims. But I have not yet found an argument that refutes either.'

that practical reason ought to become more impersonal, it is not entirely transparent whether these implications must be understood in an ‘extreme’ or ‘moderate’ way. As I hope to make clear in the rest of this thesis, much rests on this distinction. Hence, in the remainder, I am going to ignore Parfit’s own somewhat awkward positioning in between these opposing claims. I shall concentrate on the extreme claim first and only turn to the moderate claim in the next chapter.

3.1 Prudence and Future Selves

I want to begin by exploring the consequences of reductionism for prudential or self-interested reasoning. To illustrate the problem, let us assume that Jack is a heavy smoker. Jack is well aware that as a result he might die twenty years from now. But why should he care? The standard reply rests on two basic arguments. First, Jack has prudential reason to care because it is *his* future. The reason for him caring is *agent-relative*. Second, his future well-being is equally his well-being. The reason why he should care is *time-neutral*. His future interests are no less important than his present interests.⁴³

There is, of course, one further assumption on which the standard reply depends: it must really be Jack who will die in agony. Now, on the Ego theory this condition may (perhaps) be met. But on the bundle theory, this is much less evident. There are two ways in which this is less evident. First, it is not inconceivable that the psychological relatedness between Jack and John will be weak. So it is unclear to Jack that *he* will suffer terribly later. Second, even if Jack and John were strongly linked, we could still deny that their psychological relatedness provides the proper agent-relative reason for prudential concern.

The extreme claim advances the latter version of the argument. It asserts that, if reductionism were true, Jack could *never* have a genuinely agent-relative reason to care

⁴³ For the view that prudential reason is time-neutral and agent-relative, see for example T. Nagel, *The Possibility of Altruism*, Chapter 8. J. Rawls, *A Theory of Justice*, pp. 293-298.

about a future person-stage of his.⁴⁴ According to this claim, there could be such a reason only if Jack and John were perfectly identical. Only this kind of identity would guarantee that he would be the same self in the appropriate, reason-giving sense. It could be said that the extremist argues for what Parfit describes as the ‘full relativity of reason’.⁴⁵ Standard prudential reason is agent-relative but neutral across time. The extreme claim argues that reasons should also be relative across time. Not only does Jack have an agent-relative reason to care for himself rather than for others. He also has reason to care for his present self rather than for the well-being of a future self. To put it differently, his future self is like another person, whose interests or preferences do not figure as a basic element in his prudential calculation.⁴⁶

What is the argument for temporal relativity? The argument ultimately revolves around the normative appreciation of relational facts. The standard view supports a very sharp and far-reaching distinction between a person on the one hand and the social relations between persons on the other. If John were a different person, his suffering would carry no direct or non-derivative weight for Jack. Nor would it be of ultimate significance, on this view, if Jack had been closely related to John. Interpersonal relatedness does not matter in the sense that it facilitates agent-relative reasons. However, if this last claim were true, why still be neutral across time? What is so special about my future person-stages in contrast with my friends or relatives or neighbours? After all, we know now that personal identity is itself merely one kind of relatedness among many. Put differently, the extremist applies the agent-relativity of reason not just to the separateness between persons, but also to the separateness between person-stages.

⁴⁴ D. Parfit, *Reasons and Persons*, pp. 307-311.

⁴⁵ *Ibid.*, pp. 140-141.

⁴⁶ For a more explicit defence of the extreme claim, see J. Stone, ‘Parfit and the Buddha: Why There Are No People’, *Philosophy and Phenomenological Research*, Volume 48, No. 3. (March, 1988), pp. 519-532.

Notice that on this view the strength of the relatedness is entirely inconsequential. Jack may be psychologically connected to John, but this is still no reason to take his suffering more seriously than the suffering of others. Granted, he may have longstanding ideals, plans and hopes, and John may have many memories of Jack's earlier experiences. Indeed, all these connections may give his life the relational coherence of a Humean bundle of experiences. But psychological relatedness, however strong, will still not suffice to give him an agent-relative reason to prevent his future misery. Thus, although Jack may be psychologically linked to John, both remain separate agents. Consequently, it is not imprudent or irrational for Jack to pay no heed to his future suffering.⁴⁷

Notice further that the extreme claim could still condemn imprudent behaviour by shifting towards moral reasons.⁴⁸ According to Parfit, the standard view of prudential reason is flawed because it is a hybrid.⁴⁹ It is agent-relative, but time-neutral. Alternatively, there are two pure theories. First, we could resort to the full relativity of reasons. This is the pure theory of rationality. Second, we could advocate the full neutrality of reasons. This is the pure theory of morality. Moral reasons are agent-neutral as well as neutral to time. According to morality, our future selves could be compared to future generations. Whereas most of us believe that we are one and the same person throughout our biological lives, and possibly even afterwards, reductionism permits that one biological body may sustain several persons. Consequently, we have to say that *someone* will suffer terribly as the result of my actions. That person will not be me, but there will be terrible suffering at some point in the future. In other words, while the extreme claim rules out an appeal to prudence, it

⁴⁷ Whiting distinguishes between a 'weak' and 'strong' version of the extreme claim. The former claims that we are *not required* by reason or rationality to care about the future. The latter claims that it is *irrational* to have concern for your future person-stages. Here, I am interested in the extreme claim in its weak version. See J. Whiting, 'Friends and Future Selves', *Philosophical Review*, Volume 95, Issue 4 (1986), pp. 547-580.

⁴⁸ D. Parfit, *Reasons and Persons*, p. 106

⁴⁹ *Ibid.*, pp. 137-148.

unlocks the door to morality. Morality suggests that Jack has reason to prevent John's agony, not because it is rational, but because he has a duty to avert terrible future experiences, regardless of whose experiences they are. If the extremist is correct, we may have a moral duty to act 'prudentially'.

3.2 Desert, Responsibility, Obligation and Past Selves

Essentially, the extreme claim asserts that the relatedness between temporal person-stages is of no non-derivative significance. From the practical point of view, the fact that person-stages are connected and/or continuous in the ways that Locke and Hume describe bears no genuine weight. That is, these relations ought to have no greater impact on our rational and moral judgements than the relations we have with other people. We have seen what this implies for prudential reason. The interests of my future person-stages do not count as reasons for me now. What, then, are the implications for backward looking concepts such as obligations of commitment, desert, responsibility and punishment?

Locke himself suggests one astonishing answer. According to him, a person cannot deserve punishment, on principle at least, for a crime committed in the past if that person has no recollection of that crime.⁵⁰ Reid suggests, even more radically that no one could *ever* be held responsible for any of his past deeds.⁵¹ I cannot, after all, deserve punishment or reward for something I did not do. Desert and responsibility, that is, can only be allocated to those entities that meet the proper identity conditions. Clearly, I may well be related to an earlier person-stage. And I may even remember doing this or that. But that cannot justify that I should now be held responsible for these actions. Stone makes the point as follows:

If I say that I feel remorse because I murdered my brother, you would be right to be puzzled. For all this comes to is that I happen to stand in certain causal

⁵⁰ J. Locke, *Essay Concerning Human Understanding*, pp. 346-347.

⁵¹ T. Reid, 'Of Mr. Locke's Account of our Personal Identity', p. 117.

relations to the man who killed my brother and no one else does, and why feel remorse over that? Why feel guilt? And if identity is merely psychological continuity it becomes impossible to account for the fact that people bear responsibility over time. 'I am the man who did the deed,' the criminal confesses, 'but all this comes to is that I have the misfortune to stand in certain causal relations to the man who did it and no one else does. I'm not to blame for this situation; I find myself this way. Punishing me for his crime is no better than punishing the son for the crimes of his father.'⁵²

Reductionism, on this extreme view, is simply incompatible with notions like responsibility and desert.

Similar implications follow for commitments and promises. A promise retains its validity over a period of time only if 'promisor' and 'promisee' retain their perfect identity.⁵³ John may be related to Jack, but this cannot bind him to a commitment or a promise originally assumed by Jack. Their psychological relatedness is, in this sense, simply irrelevant. Jack may have promised to meet you for lunch. But that promise cannot obligate John. Jack's promise is not merely a past commitment; it is the commitment of another person.

Desert, responsibility and obligation are backward looking in the sense that they presuppose a view on which present and future selves could be bound by the actions, intentions and commitments of an earlier self. Past intentions, promises and actions, that is, must carry normative weight such that I can later be said to deserve, owe, be obligated, and so forth. The extreme claim, however, denies that personal identity has these normative aspects. Psychologically relatedness has no normative upshots.

3.3 Justice and the Separateness of Persons

There is yet another way in which reductionism might have important implications. Parfit argues that it lends credibility to impersonal versions of utilitarianism. It is in arguing for this position that Parfit's own view becomes almost indistinguishable from

⁵² J. Stone, 'Parfit and the Buddha: Why There Are No People', p. 529.

⁵³ D. Parfit, *Reasons and Persons*, p. 326.

the extreme claim.⁵⁴ Rights-based theories of justice, of the kind advanced by Rawls, Nozick and Dworkin, commonly claim overriding importance for persons and personal identity.⁵⁵ This claim is often described in terms of the separateness of persons. The appeal to this separateness revolves around what Parfit calls the ‘compensation claim’. This states that the benefits of one person cannot compensate the burdens of another. So we cannot justify imposing certain costs on Ralph, merely by claiming that this will greatly benefit Paul. To do so would be unfair. Extreme cases aside, this is true even if the benefits outweigh the costs. It is for this reason that, apart from an increase in benefits, these theories are also interested in finding just distributions of burdens and benefits. Thus, the notion of the person puts constraints on consequentialist morality. About the tradition of classical utilitarianism, Rawls remarks that it ‘does not take seriously the distinction between persons’, and mistakenly adopts ‘for society as a whole the principle of rational choice for one man.’⁵⁶ Utilitarianism, that is, ignores the crucial distinction between social justice and rational prudence. As he puts the criticism: ‘if we assume that the correct regulative principle for anything depends on the nature of that thing, and that the plurality of distinct persons with separate ends is an *essential* feature of human societies, we should not expect the principles of social choice to be utilitarian.’⁵⁷ And Nagel writes: ‘To sacrifice one individual life for another, or one individual’s happiness for another’s is very different from sacrificing one gratification for another within a *single* life.’⁵⁸

Notice that Rawls and Nagel assume compensation within one life to be fairly unproblematic. That is, a burden for someone *now* could be compensated if benefits would accrue to that person *later*. We generally do not think it wrong to incur certain

⁵⁴ Ibid., pp. 329-345. See also B. Schultz, ‘Persons, Selves, and Utilitarianism’, *Ethics*, Volume 96, No. 4. (July, 1986), pp. 721-745; D. Parfit, ‘Comments’, *Ethics*, Volume 96, No. 4. (July 1986), pp. 832-872, at pp. 837-843.

⁵⁵ I concentrate here on J. Rawls, *A Theory of Justice* (Oxford: Oxford University Press, 1992), especially pp. 22-33. As we already saw, a similar view is found in the work of Dworkin and Nozick.

⁵⁶ J. Rawls, *A Theory of Justice*, pp. 26-27.

⁵⁷ Ibid., p. 29, my emphasis.

⁵⁸ T. Nagel, *The Possibility of Altruism*, p. 138.

burdens presently if this will pay off later on in life. Trade-offs over time are not objectionable, then, as long as they occur within the life of one person. Thus, it is evident how much rights-based theories rely on sharp boundaries between persons. Within the unity of a life compensation and trade-offs are permitted. Here, we can just maximise. But as soon as costs and benefits are assigned to separate persons, maximisation should be abandoned.

However, on the reductionist account of personal identity, there are no fixed and clear boundaries. According to Parfit, reductionism has a dual effect on theories of just distribution. First, it leads to an extension of the *scope* of justice. That is, benefits and burdens ought also to be distributed fairly between the temporal person-stages that make up our lives. As Parfit observes, the question ‘When?’ should be compared to the question ‘Who?’⁵⁹ As a result, the maximising principle should be constrained not less, but further. Incurring a present cost in order to reap a later benefit is precisely comparable to the case of Paul and Ralph. Future benefits do not make up for present costs because my future self and I are not strictly the same person. We may be psychologically related, but compensation still constitutes an injustice. After all, Paul may also be psychologically related to Ralph, but this fact does not justify compensation either.

However, Parfit goes on to show that the principles of fairness and equality lose *weight* at the same time. In fact, the entire notion of distribute justice becomes blurry at this point. Over what should burdens and benefits still be distributed? Why is it bad that some bundles of experiences are miserable and others happy? Surely, what is more important than the distribution of misery and happiness over such bundles, is that there are as few as possible miserable experiences. The problem, Parfit suggests, is that the units over which we distribute sink to the micro-level of temporally discrete experiences. But at this level, justice as fairness is simply on par with utilitarianism.⁶⁰

⁵⁹ D. Parfit, *Reasons and Persons*, p. 340.

⁶⁰ *Ibid.*, p. 344.

Rawls argues that classical utilitarianism ‘extends to society the principle of choice of one man’ and that it does not recognise the separateness of persons as ‘an *essential* feature of human societies’. Parfit points out that Rawls is making a similar mistake. He extends to man the principle of rational choice for one temporally discrete person-stage, and he fails to see that the separateness of *person-stages* is an essential feature of human life.

The bottom line can be expressed in terms of the analogy between individuals and groups. The extreme claim maintains that individuals are like associations not just from the ontological point of view, but also from the point of view of justice. Justice as fairness presupposes that the well-being of persons has non-derivative and reason-giving value. But Parfit claims that persons are of no greater moral importance than nations. From the point of view of justice, the boundaries between people are like boundaries between countries. They delineate nothing of ultimate significance. Individualists agree when he claims:

If there is nothing more to a nation than its citizens, it is less plausible to regard the nation as itself the primary object of duties, or possessor of rights. It is more plausible to focus upon the citizens, and to regard them less as citizen, more as people. We may therefore, on this view, think a person's nationality less morally important.⁶¹

However, because persons are really only groupings of experiences, we should focus more on ‘experiences themselves’ rather than the ‘subject of experiences’. So, he goes on to argue analogously:

It becomes more plausible, when thinking morally, to focus less upon the person, the subject of experience, and instead to focus more upon the experiences themselves. It becomes more plausible to claim that, just as we are right to ignore whether people come from the same or different nations, we are right to ignore whether experiences come within the same or different lives.⁶²

⁶¹ Ibid., p. 340.

⁶² Ibid., p. 341.

Of course, Parfit is always careful to couch his argument in terms of relative plausibility, but on a more bold statement, we might say that individuals, like nations, are no more than the receptacles of what really matters. In fact, moral individualists are exactly like extremists when considering relatives, friends and nations. They do not think that such ties and relations are themselves reason-giving. To use Godwin's famous words, there is no magic in the pronoun 'my'.⁶³ But the extremist's point against ethical individualism is precisely that persons, too, are associations.

Conclusion

This chapter argues for the inversion of the analogy between individual and group. We no longer believe in Gierkean collective Egos. Reductionism about personal identity suggests that we should also stop believing in the Self. On this basis, individuals and groups are once again on a level pegging. Associations are not super-individuals, but individuals are miniature associations.

This analogy has two different aspects. The bundle theory of personal identity provides the ontological side of the argument. It is a common view that groups consist merely in a plurality of individuals who are related or associated in a specific sense. According to the bundle theory, something similar applies to the individual itself. Personal identity is based on the psychological relatedness between intentional episodes or person-stages. A continuous person is an association *between* beliefs, desires and intentions at different points in time, not some 'further fact' that exists *underneath* them.

The practical side of the argument claims that, as a consequence, our theories of rationality and morality need serious revision. According to the extreme claim, reasons should become more impersonal. We should regard persons more like ethical

⁶³ W. Godwin, *An Enquiry Concerning Political Justice*, in: Philp, M., (ed.), *Political and Philosophical Writings of William Godwin, Volume 3* (London: William Pickering, 1993), Book II, Chapter II, p. 50.

individualists regard nations and other associations. Ethical individualists believe that the value of nations and states derives from the extent to which they contribute to individual well-being. The extreme claim holds that the value of personal identity is similarly derivative. Like associations, the good of an individual is not valuable in a fundamental sense, but only insofar as this contributes to the well-being of person-stages. Moreover, the extremist claims that we should reject notions of desert, responsibility and obligation.

The analogy between groups and association, then, appears to hold up in spite of everything. But, it will be objected, does it actually strengthen the case for corporate personality? For although individuals are like groups, it is also obvious that this is a Pyrrhic victory. The problem is that we have levelled down rather than up. We may only have shown that the idea of a person should be expunged from rationality and morality altogether. Chapter II aims to circumvent this objection. What that chapter looks for is an argument that could demonstrate that continuous persons, rather than temporally discrete person-stages, should remain the basic unit in practical analysis, *in spite of* the fact that personal identity is relational. Once this argument is up and running, the remainder of this thesis is no more than an attempt to extend its scope to interpersonal associations.

CHAPTER II

INTEGRITY

Like our conception of *eudaemonia*, unified agency is an achievement – in fact, the same achievement, seen from a different point of view. The process resembles the day-to-day activities of a sculptor I know. He begins work with a pile of found metal objects – rebar, cotter pins, washers, steel plate, and so on – which he pieces together to form larger units: the head of the hoe is joined to a length of bent reinforcement rod, and is now a hand. Similarly, we start our practical lives with haphazard collections of desires so poor in content as to amount to no more than reflexes – some innate, some conditioned, and some supplied by adults around us. Pressured by experience to resolve practical conflicts, we weld disparate desires into larger and more structured practical judgments.

Elijah Millgram¹

Introduction

According to one objection, corporate realism simply had to fail. It had to fail for the reason that it was ultimately based on an untenable metaphysical theory of collective Egos. As this theory no longer has credibility, corporate personality is defensible only on the basis of juristic fiction, leaving the domain of moral personality exclusively to the individual.

However, it is now evident that the issue is more complicated. In Chapter I it emerged that the notion of the individual person faces a similar objection. There is a view, we saw, that claims that ethical individualism is based on a metaphysical theory of personal identity over time. This theory grounds personal identity on the bedrock of the Self that exists separately from our brain and our thoughts. But this theory has also lost much of its credibility. Consequently, just as little stands in the way of discarding ethical individualism.

¹ 'Incommensurability and Practical Reasoning', in: R. Chang (ed.) *Incommensurability, Incomparability, and Practical Reason* (Cambridge: Harvard University Press, 1997), pp. 151-169, at pp. 162-163.

Such a move would rest on two claims:

(P1) *Psychological reductionism* claims that selves are temporally punctuated and not grounded in the substratum of a Self. On this view, personal identity over time consists in the degree of relatedness between mental events, such as thoughts, intentions and beliefs, occurring at different times.

(P2) The *extreme claim* asserts something about the normative aspects of personal identity, conceived as such a temporal relationship. It argues that temporally extended persons should no longer be a basic unit of analysis in morality and rationality. All things equal, the fact that two person-stages are in some psychological sense connected is of no greater reason-giving importance than the fact that two individuals are so related. From the standpoint of morality and rationality we ought to assign fundamental weight only to the stages of a person. That is, every intentional episode should be seen as a discrete agent, with basic interests and rights in and of itself.

If true, these claims leave the realm of real personality completely empty of inhabitants and thus strike at the very heart of liberalism as we find it in the political theories of Locke, Kant and Rawls, Nozick and Dworkin. It suggests a radically impersonal form of utilitarianism, according to which temporally isolated person-stages form the only unit of well-being and responsibility.

Together with these theorists, I believe the latter view to be false. But in order to be able to claim this, we need to reconfigure our picture of what it means to be a person. This means that we have to reject either (P1) or (P2). I want to argue in this chapter that we should reject the latter and accept its counterpart:

(P3) The *moderate claim* maintains that even though personal identity should be conceived of as a relationship, the person is still the basic unit or currency in morality and rationality. All other things being equal, the fact that two person-stages are in some specified sense connected is just about as significant as the unity of a person on the non-reductionist view. Thus, we should give normative priority to the bundle or relatedness – what I below describe as integrity – that makes us into continuous persons, rather than the interests and desires of the temporal stages of a person.

Why accept P3 rather than P2? The key to this question is found in two important arguments.

The first argument is what I generally refer to as the *practice theory*. This theory captures something akin to what Rawls proposed in his essay ‘The Independence of Moral Theory’.² Instead of grounding normative theory in a conception of the person and personal identity, we have to begin by assuming that certain practices are a basic feature of human life, at least as we know it. And from these practices we then reason towards a certain general and normative understanding of the person and personal identity. In sum, the practice theory argues that we need to accept the moderate claim because certain basic intuitions, judgements and practices would not be intelligible otherwise.

Second, we need to extend Parfit’s account of R-relatedness with a normative dimension. On the bundle theory, personal identity consists in the *psychological* relatedness between person-events, where each of these events is just an experience being had. We need a better and more specific theory of what we mean by this relationship. Basically, we need a conception of personal identity that defines this relationship in *normative* terms. Among other things, we need a conception on which person-stages are connected in such a way that prudence, responsibility and commitments are possible. I shall refer to this normative dimension of identity as *integrity*.

In this chapter, I want to say more about these two claims. However, it is important to warn that in pressing these claims my ulterior aim is seditious. Ultimately, I want to show that the moderate claim can be applied, not just to *intrapersonal* but also to *interpersonal* relatedness. My aim is covert, so to speak. It is to show that the practice and integrity-based approach restore the possibility of corporate personality.

This is a concern for later chapters. Here, I proceed by introducing the moderate claim in Section 1. Section 2 outlines the practice theory of personal identity. In Section 3, I discuss a number of objections. Section 4 takes up the idea of integrity. Then, in

² J. Rawls, ‘The Independence of Moral Theory’.

Section 5, I suggest that integrity is a virtue separate from morality. Finally, Section 6 puts the idea of personal integrity back into a temporal context.

1. The Moderate Claim

The extremist does not turn a blind eye to the fact that most of our lives are psychologically connected in the way that Locke and Hume describe. However, the issue turns entirely on how this relational fact is appraised. On the one hand, the extreme claim denies that R-relatedness is reason-giving in various ways. So the connectedness between person-stages does not carry with it obligation, desert and guilt, justice does not allow for trade-offs, and future preferences do not count as present reasons. The separate parts of a person are conceived to be separate agents. On the other hand, the moderate affirms that psychological connectedness has precisely this kind of moral and rational significance. To put it differently, the moderate wants to restore the temporal unity of agency by claiming that personal identity refers to a kind of relatedness that is reason-giving in the various ways that the extremist denies.

What, for example, does the moderate claim imply for our standard conception of prudential reason? Take again the case of Jack, a heavy smoker, and his future self, John, who dies in agony as the direct consequence. Assume that Jack and John are psychologically connected. Does Jack have reason to take John's interests into consideration? Essentially, the moderate argues that prudential concern is *inherent* in the idea of personal identity. The idea is not so much that psychological relatedness justifies prudential concern – as something external to and detachable from the relation itself – but rather that prudence is *constitutive* of personal identity. The moderate claim has been rightly compared to friendship in this regard.³ It is a common view about friendship that it implies mutual concern for each other's interests. Again, the point is

³ J. Whiting, 'Friends and Future Selves'.

not that friendship justifies concern, but that the two cannot be analytically distinguished. To be a friend simply *is* to care about that person's well-being in a particular way. In other words, the concept of friendship is defined in terms of mutual concern. According to the moderate, the connectedness between present and future selves should be seen in the same light. To be related as one continuous person simply *is* to care about your future selves.

The moderate claim has a similar impact on backward looking concepts such as desert, responsibility and obligation. The moderate argues that personal identity should be understood to carry with it responsibility for past crimes. While personal identity is a relation, it is the kind of relation in which mutual accountability is inherent. Consider the following case. I hatch a plan to rob the National Gallery, which I then proceed to carry out in a step-by-step manner. Altogether the plan comprises many separate sub-plans and many separate actions. I start by making all the required preparations. I buy equipment, arrange for a get-away-car, and so forth. On the night of the robbery, I switch off the alarm, sneak in through the back, stun the guards and make off with a painting. Now, if the extremist were correct, we would have to break down this sequence in temporally discrete actions, and then allot moral responsibility for these actions to each of the person-stages. One stage is responsible for illegal entry, another for stunning the guards, and so on. We could, possibly, judge the wrongness of the event by adding up the wrongness of each action, but even if this produced the right results, there is no present agent who could be held responsible. The moderate, in contrast, emphasises that personal identity is precisely the kind of relationship of which it is constitutive that the members assume responsibility for each other. My person-stages, that is, intend to achieve something together, and they embrace each other's actions as their own. They are collectively accountable.⁴

⁴ The notion of collective responsibility is, of course, the subject of ongoing controversy. See L. May and S. Hoffman (eds.), *Collective Responsibility: Five Decades of Debate in Theoretical and Applied Ethics* (Savage: Rowman & Littlefield Publishers, 1991). However, the bundle theory suggests that, if

By now the gist should be sufficiently clear.⁵ The moderate claim retains person-based ethics, while at the same time it embraces the reductionist truth about the identity of persons. Reductionism about personal identity fragments our agency in intentional episodes. But the moderate claim re-establishes the unity of agency by arguing that these episodes share responsibility for their actions and put each other under obligations. To put it differently, the moderate argues that identity is a relation with normative standing.

2. The Practice Theory

The attempt to undermine person-based theories of rationality and morality by advancing a reductionist theory of personal identity could be resisted, I argue, by putting more emphasis on the moderate claim. We do not need a non-reductionist conception of personal identity; we just need to argue that persons continue to matter morally and rationally, even though their temporal identity is relational. The knack is to deepen the psychological relatedness between person-stages with a normative dimension. Section 4, 5 and 6 present a more detailed account of this dimension of personal identity. But I first want to look at another question.

Why should we be moderates at any rate? I already mentioned that Parfit thinks that the extreme claim and the moderate claim are *both* defensible. This dithering, I believe, stems from the fact that no metaphysical argument could in principle decide this issue. Ontology tells us what personal identity *is*. It clearly does not tell us how personal identity should be *appraised*. For this we need a different sort of argument, a kind of

responsibility is possible at all, the difference between individual and collective responsibility is much less profound than often believed.

⁵ Similar claims can be made about commitments to third parties and the importance of the separateness of persons. That said, it is probably more important to mention here that on the moderate claim the weight of prudential reasons and of obligations, desert and responsibilities may be mediated by the strength of the connectedness. For example, it may be rational to discount my future interests at a rate proportional to the degree to which my psychological connectedness weakens through time. I will come back to this issue in Chapters V and VI.

argument, I claim, that is in itself normative. I proceed by unpacking this argument below.

Notice first that Parfit would reject such a practice-based argument.⁶ Buried in Parfit's writings on personal identity lies a particular view about the order in which normative theory should be conducted. The presupposition is that normative theory should in the end be grounded in some account of what persons really are. This account of the self will ultimately reveal the truth about rationality and morality. Parfit, then, seems to believe that we need to begin by ironing out the issue of personal identity. For only with this out of the way could we pass a judgement on human practices.

What I shall refer to as the practice theory departs from this method. The limitation of Parfit's approach is precisely that it cannot arbitrate between the extreme claim and the moderate claim. The practice theory proposes that we therefore reverse this procedure. It begins by looking at certain basic human practices and judgements and intuitions and then tries to manufacture a theoretical conception of personal identity that achieves a decent fit with these data. For example, we worry about going to the dentist, we get nervous about job-interviews and many people agonise over growing old. We also tend to think that promises create future obligations. It may of course be true that 'promisor' and 'promisee' are merely earlier selves, but we still think, all things being equal, that promises have an obligating force. We are, moreover, all engaged in numerous kinds of enduring social relationships and networks. Friendship, trust, citizenship and contracts are difficult to imagine without assuming the temporal continuity of a person.⁷ The vows we make, the contracts we sign, the love for our children, the debts that need to be repaid, all these everyday institutions and practices presuppose a sense of personal continuity. Characteristically, when we fail to live up to

⁶ Cf. T. Chappell, 'Reductionism about Persons; And What Matters', *Proceedings of the Aristotelian Society*, Vol. XCVII (1997-1998), pp. 41-57, p. 47.

⁷ See S. R. Wolf, 'Self-Interest and Interest in Selves', *Ethics*, 96 (1986), pp. 704-720; R. M. Adams, 'Should Ethics Be More Impersonal?', in: J. Dancy (ed.), *Reading Parfit* (Oxford: Blackwell, 1997) pp. 251-289, in particular pp. 263-273.

our social responsibilities, we cannot simply get ourselves off the hook by saying: ‘It was not I, but an earlier self!’

While Parfit clearly wants to review those ethical practices on the basis of an antecedent theory of personal identity, the practice theory evaluates conceptions of personal identity on the basis of these and other practices and intuitions. Moreover, if we proceed along this order, it becomes evident that we really have no choice but to accept the moderate claim. The attractiveness of the extreme claim quickly evaporates when we consider, as Wolf puts it succinctly, ‘the way in which our identification of and our concern for ourselves and each other as *persons* essentially contributes to, if you’ll pardon the expression, our form of life.’⁸

The idea that we could rebut the extremist argument by inverting the relation between personal identity and normative theory emerges in bits and pieces from a number of diverse practical angles, each raising what is at a deeper level the same objection. Rather than providing the grounding premise of a theory of rationality and morality, so that objection goes, a conception of a person is a conclusion of normative theorising. And this means that theories of the person should be validated in what is commonly known as reflective equilibrium or some form thereof.⁹ Like other principles, then, the extreme and moderate claims are to be tested against our well-considered judgements and intuitions on normative matters in general. I shall now briefly discuss three more specific versions of this objection to Parfit’s extreme thesis and then focus on a number of counter objections.

First, Korsgaard and Blackburn present a broadly Kantian version of the practice theory.¹⁰ We could describe it as the argument from action ownership. This argument

⁸ S. R. Wolf, ‘Self-Interest and Interest in Selves’, p. 708.

⁹ For the notion of reflective equilibrium, see J. Rawls, *A Theory of Justice*, pp. 46-53; R. Dworkin, *Taking Rights Seriously*, pp. 159-168.

¹⁰ C. M. Korsgaard, ‘Personal Identity and the Unity of Agency: A Kantian Response to Parfit’, *Philosophy and Public Affairs*, Volume 18 (1989), pp. 101-132, especially pp. 119-121; S. Blackburn, ‘Has Kant Refuted Parfit?’, in: J. Dancy (ed.), *Reading Parfit* (Oxford: Blackwell Publishers, 1997), pp. 180-201.

appeals to Kant's distinction between practical and theoretical reason. Briefly, the former is concerned with the *justification* of action, while the latter essentially aims at the *explanation* and *prediction* of behaviour. Practical reason is essentially locked into the first-person-point-of-view of the agent who needs to decide what to do. Theoretical reason requires the external standpoint of the objective observer, who studies the world but is not part of it.

Now, according to this strand of the argument, we may well build a case for reductionism from the disengaged perspective. But once we assume the perspective of practical reason, we are forced to conceive of the issue in a different light. Essentially, we have to view ourselves as deliberators. And this means that we have to conceive of ourselves not passively as 'experiences being had', but rather as the free choosers and therefore the owners of our words and actions. Korsgaard writes:

... from the practical point of view our relationship to our actions and choices is essentially *authorial*: from it, we view them as *our own*. I believe that when we think about the way in which our own lives matter to us personally, we think of ourselves in this way. We think of living our lives, and even of having our experiences, as something that we *do*.¹¹

And Blackburn similarly claims:

When I deliberate, I wonder what to think or do: I do not try to predict, passively from the outside, what I shall do. When I judge, I wonder what to say about something: I do not try to predict, passively and from the outside, what I shall end up saying about it. These activities proceed under the discipline of reason, and hence are subject to norms, and they involve myself as an agent, not the latest stage of a bundle of separately individuated psychological states. When I think, I as it were take charge of the bundle, and it is only from the outside perspective that this is to be thought of as the passive arrival of more states for it.¹²

Clearly, the extreme claim is fundamentally at odds with the deliberative point of view of agency. If the extremist were correct, we could never authorise and guide action at all. The very idea of practical deliberation and action ownership appears to rest on the a-

¹¹ C. M. Korsgaard, 'Personal Identity and the Unity of Agency: A Kantian Response to Parfit', p. 121.

¹² S. Blackburn, 'Has Kant Refuted Parfit?', in: J. Dancy (ed.), *Reading Parfit*, p. 192.

priori conception of temporally extended agency, that is, a view on which person-stages are rationally and morally integrated. As Korsgaard and Blackburn point out, it is ultimately inconsistent to advance a conception of personal identity that contradicts our capacity for deliberation.¹³ Assuming that the extremist still wants to say something about rationality and morality, he cannot at the same time deny what is the pre-condition of practical reason.

Second, Taylor, MacIntyre and Williams advance a different version of what I think is at bottom the same objection.¹⁴ Their argument is that we can only come to grips with the issue of personal identity from the vantage point of the good. Human beings, so the claim goes, inevitably have to face up to the question of the goodness of their lives taken as a whole. Our self-understandings take shape

... in a certain space of questions, through certain constitutive concerns. The questions or concerns touch on the nature of the good that I orient myself by and on the way I am placed in relation to it. But then what counts as a unit will be defined by the scope of the concern, by just what is in question. And what is in question is, generally and characteristically, the shape of my life *as a whole*. It is not something up for arbitrary determination.¹⁵

What makes a life worth living? And how does my life measure up to this standard? According to Taylor, these questions have a certain inevitability, if not universally, then at least in our culture. We inhabit 'moral space'. And because we need orientation and direction in this space, it is equally inevitable that we require conceptions of the good, plans or 'ground-projects' (as Williams calls them). We all have and need such conceptions, for ground-projects are 'the condition of my existence, in the sense that

¹³ P. Pettit, *A Theory of Freedom* (Cambridge: Polity Press, 2001), in particular pp. 82-85, similarly claims that 'discursive control' is constitutive of the idea of a person. At p. 83 he claims: 'What is said to make two people at different times the same person is that they are connected in such a way – whatever it is – that under existing discursive practice the later can be held responsible for and to the earlier.' See also C. Rovane, 'Self-Reference: The Radicalization of Locke', *The Journal of Philosophy*, Vol. 90, Issue 2, (1993), pp. 73-97; *The Bounds of Agency: An Essay in Revisionary Metaphysics*.

¹⁴ C. Taylor, *Sources of the Self: The Making of the Modern Identity* (Cambridge: Cambridge University Press, 1989), in particular Chapter 2; A. MacIntyre, *After Virtue: A Study in Moral Theory* (London: Duckworth, 2001), tenth impression, especially Chapter 15; B. Williams, 'Persons, Character and Morality', in A. O. Rorty (ed.), *The Identities of Persons* (Berkeley: University of California Press, 1976), pp. 197-216.

¹⁵ C. Taylor, *Sources of the Self*, p. 50.

unless I am propelled forward by the conatus of desire, project and interest, it is unclear why I should go on at all'.¹⁶

The extreme claim fails to come to terms with this basic feature of human life. It fails to grasp that human agents need an anchoring in life-plans, which will also provide them with the kind of diachronic unity that gives us our identity.¹⁷ As Williams notes: 'the problems that face Utilitarianism about agency can arise with any agent whose projects stretch over enough time, and are sufficiently grounded in character, to be in any substantial sense *his* projects'.¹⁸ To put it differently, the entire notion of the goodness of a life as whole is inconsistent with the extreme claim. According to the extremist, it does not matter that a life is good. What matters is that there are good *experiences*. The whole question of what makes a *life* go well could not even arise.

We could describe this last objection to the extreme claim as the argument from the good. Rawls offers a third and justice-based version of the same objection.¹⁹ His 'political' conception of the person is tailored to fit within a wider, democratic tradition that conceives of society as a fair system of co-operation over time. This democratic conception, he argues, is deeply embedded in the public culture and the political practices of Western societies. Rather than positioning itself outside history, Rawls's theory of justice simply aims to construct the principles and ideas that are implicit in this culture and practice. Essentially, then, the claim is that the principles of justice organise what already are 'our most firmly held convictions', for example, that slavery is wrong. This practice-based approach also shapes the Rawlsian conception of the person.

¹⁶ B. Williams, 'Persons Character and Morality', p. 208.

¹⁷ Robert Nozick would appear to have agreed: 'A person's shaping his life in accordance with some overall plan is his way of giving meaning to his life; only a being with the capacity to so shape his life can have or strive for meaningful life.' See his *Anarchy, State, and Utopia* (New York: Basic Books, 1974), p. 50.

¹⁸ B. Williams, 'Persons, Character and Morality', p. 202.

¹⁹ J. Rawls, 'The Independence of Moral Theory'; 'Justice as Fairness: Political not Metaphysical', *Philosophy and Public Affairs*, Vol. 14 (1985), pp. 223-251; *Political Liberalism* (New York: Columbia University Press, 1993), pp. 18-20, 29-35.

There are, of course, many aspects of human nature that can be singled out as especially significant, depending on our point of view. This is witnessed by such expressions as ‘homo politicus’ and ‘homo oeconomicus’, ‘homo ludens’ and ‘homo faber’. Since our account of justice as fairness begins with the idea that society is to be conceived as a fair system of cooperation over time between generations, we adopt a conception of the person to go with this idea.²⁰

Thus, since Rawls locates himself within a democratic practice, he settles on a view of the person that allows us to fully take part in this practice. And this means, among other things, that persons have or ought to have ‘a capacity for a sense of justice’ and ‘a capacity for a conception of the good’. As we have already seen, it also means that the separateness of persons retains its significance. There is, then, no place for something like the extreme claim. That claim is so radically at odds with our political and democratic practices that it could hardly be taken seriously.

3. Three Objections to the Practice Theory

The moderate claim allows us to still engage in longstanding relationships, it validates a view in which our good on the long term importantly figures, and it further feeds into a widespread sense of justice and fairness. As opposed to this, it is far from evident what the ethical attractions of the extreme claim are supposed to be. There is the vague promise of a highly impersonal brand of utilitarianism. But the extremist, I believe, is hard pressed to persuade us that his conception of personal identity has any practical appeal at all. The problem is not just that the extreme claim looks distinctly nightmarish, it is also incredibly remote from how we actually live and understand our lives, so remote that that we find it hard even to *imagine* what it means to be a mere part of a person, to have no past and future at all. While he says little about this, Parfit confesses to the following:

Is the truth depressing? Some may find it so. But I find it liberating, and consoling. When I believed that my existence was such a further fact, I seemed

²⁰ J. Rawls, *Political Liberalism*, p. 18.

imprisoned in myself. My life seemed like a glass tunnel, through which I was moving faster every year, and at the end of which there was darkness. When I changed my view, the walls of my glass tunnel disappeared. I now live in the open air. There is still a difference between my life and the lives of other people. But the difference is less. Other people are closer. I am less concerned about the rest of my own life, and more concerned about the lives of others.²¹

But these advantages seem also available to the moderate. In fact, if the extremist were right, it would be difficult to understand why I should care about anyone at all, my own future self or some other stage of a person.²²

That said, the main thrust for the extreme claim, in all fairness, should probably not be sought in its intuitive appeal, but rather in a number of objections to the practice theory as such.

3.1 *Is Personal Identity Bifurcated?*

To begin with, Korsgaard and Blackburn seem sympathetic to the view that theoretical and practical reason provide us with two irreducible standpoints that, while obviously conflicting, have their own autonomous validity.²³ On their view, there are two concepts of personal identity: one normative and one theoretical. While they concede that Locke, Hume and Parfit may well be right about personal identity at one level, they go on to point out that their theories have no practical consequences. From the practical point of view, the idea of temporally punctuated agents is ruled out on the a priori basis of what Blackburn describes as a 'unity reaction'. The temporal unity of agency *cannot* be questioned. But if this claim is correct, then it seems we would also have to also accept

²¹ D. Parfit, *Reasons and Persons*, p. 281. For a critical discussion of the ethical attractiveness of Parfit's reductionism, see also T. Chappell, 'Reductionism about Persons; And What Matters'.

²² Cf. S. Scheffler, 'Ethics, Personal Identity, and Ideals of the Person', *Canadian Journal of Philosophy*, Vol. XII, No. 2 (June 1982), pp. 229-246.

²³ If I understand him correctly, Taylor, in his book *Sources of the Self* (Chapter 2), argues that it is on the whole wrong to conceive of personal identity from the theoretical viewpoint. Rawls, on the other hand, argues for a view on which both standpoints are reconcilable. See J. Rawls, 'The Independence of Moral Theory', p. 296, and more recently *Justice as Fairness: A Restatement*, E. Kelly (ed.) (Cambridge MA: The Belknap Press of Harvard University Press, 2000), p. 19.

a deep and, it may be objected, unhappy bifurcation between two perspectives of reason and two concepts of personal identity.

I think, however, that we could avoid paying this price. Of course, Parfit too avoids this separation of domains. But in his view this means that we have to customise the normative to the disengaged perspective. The latter demonstrates that there is no personal identity, and so he argues that we should revise our views of rationality and morality accordingly. More precisely, Parfit claims that rationality and morality rely on the metaphysical illusion of a Self from which we should rid ourselves. Difficult though he admits this to be, we ought at least to try to concern ourselves more with temporally discrete utility and less with persons per se. However, the problem with Parfit is that the disengaged view of personal identity only has consequences for normative theory when funnelled through either the extreme or the moderate claim. And these claims are not themselves part of the disengaged truth about personal identity. There is no disengaged reason to support either of these claims.

What I am suggesting here is that these funnelling claims should be judged on the basis of the practice theory. So we do not have to deny that identity is relational. What we need to deny is the extreme claim. A bifurcation between deliberative and theoretical perspectives is thus prevented. To recap, Locke, Hume and Parfit claim that persons are temporal bundles of psychological events. From the normative standpoint, and by appealing to the importance of deliberation, prudence, obligation and justice in our judgements, we should then claim that these bundles have reason-giving weight in appropriate ways. Both claims are entirely consistent. We do not have to presuppose a 'reaction of unity' that is somehow inconsistent with what we think is true when we look at the issue from the disengaged standpoint of theoretical reason.

3.2 Rule and Indirect Utilitarian Justifications for Person-based Ethics

It may next be objected to all this that a rule-consequentialist argument could be wheeled in to redeem the extreme claim. After all, Parfit only wants to argue that personal identity has in itself no *ultimate* value. Presumably, we should still attach some weight to Relation R. The idea is rather that we should only do so if and insofar as this turns out to be helpful, on the whole, to the realisation of the overall utilitarian goal, which in his case is to maximise the temporally discrete well-being of human beings. Thus, as a rule, it could still make sense to hold people to their word, to worry specifically about our own future, to punish the guilty for their crimes and to distribute equally over bundles rather than person-stages. It could still be that, in service of maximisation, we should treat people as more or less continuous persons. It could even make sense to promote the kind of non-instrumental attitudes and dispositions that tend to bolster compliance to these rules, for example, the conviction that persons should be treated with basic respect.

But there are well-known problems with this view. Most importantly, while it is true that this *could* be the case, it is an unacceptable simplification to suggest that following these rules, conventions and dispositions *will* indeed always facilitate the utilitarian goal. Whether it will or not, is ultimately an empirical question. Following a rule may be maximising nine times out of ten. But what do we do the tenth time? If we want to remain loyal to the initial utilitarian thought, then we would have to reform and fine-tune that rule to the circumstances. David Lyons has pointed out that this means that utilitarianism must spell out not just *what* rules are useful, but also *when* they are usefully complied with and by *whom*.²⁴ And once we start accommodating exceptional and special cases within the rule, then the level of complication simply becomes too high for moral theory to generate clear results. If I understand Lyons correctly, his objection is that the rule-consequentialist principle requires a level of empirical

²⁴ D. Lyons, 'The Moral Opacity of Utilitarianism', in: B. Hooker, E. Mason, and D. E. Miller (eds.) *Morality, Rules, and Consequences: A Critical Reader* (Edinburgh: Edinburgh University Press, 2000), pp. 105-120.

information that we cannot possibly possess. In the absence of this information, rule-utilitarianism is merely indeterminate.

Moreover, rule-utilitarianism may or may not describe what morality demands, but it clearly does not offer a motivation for human agents to follow the rules. After all, we think a man punishable, not because punishment promotes utility in general, but because of his culpability. And when I break my word, you will feel personally hurt and angry. There could be indirect utilitarian justifications for the inculcation of non-utilitarian kinds of attitude that would motivate rule-following. But if this is the solution, utilitarianism is going to include the intuitively awkward idea of the state having to publicly promote the discourse of person-based ethics, knowing full well that this discourse is false. This kind of duplicity, which has been described as Government House utilitarianism, is hard to reconcile with an open and democratic society.²⁵

3.3 Pluralism and the Practice Theory

Finally, Parfit's extremist may object that the practice theory begs the question in assuming what needs to be explained. Certainly, it is true that most of us feel that selves matter – our own selves in particular. But the whole point of a theory, it will be claimed, is to investigate and challenge these pre-analytical beliefs on the basis of a more or less objective criterion. The point is that we ought to become less selfish, less preoccupied with persons, and more with the experiences of person-parts. Why have long-term projects and relations and responsibilities? Why distribute? Why engage in deliberative action?

The best response to this objection may simply be to deny that there are foundational answers to these questions to begin with. I certainly do not possess them, but nor do Parfit and the extremist. As Parfit admits, there appears to be no disengaged reason to

²⁵ See J. J. C. Smart and B. Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1973), pp. 138-140; A. Sen and B. Williams, 'Introduction: Utilitarianism and Beyond', in: A. Sen and B. Williams (eds.) *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1994), pp. 15-16.

prefer the extreme claim to the moderate claim. And this is the point where the objection gets stuck. If we abandon the foundational project, normative theory is commonly understood to operate on the basis of a different method. According to this method, normative theories and conceptions need to be brought on the same wavelength with our considered judgements and intuitions on the matter. That is to say, our theories need to be born out in a version of reflective equilibrium. Of course, the notion of a reflective equilibrium has been criticised for being inherently conservative.²⁶ And indeed, there would be little point if a theory perfectly reproduced all our intuitions and judgements. But an appropriate understanding of reflective equilibrium entails that our theories should simplify, re-organise and abridge the data to make them cohere at the deeper and more general level of principle. This will inevitably mean that some of the pre-analytical intuitions will have to be rejected. That said, conceptions of the good life, long-term relationships and social responsibilities form so overpowering a part of our lives that the moderate claim is likely to provide the better theory on any understanding of reflective equilibrium.

To be sure, Parfit goes out of his way to show that it is at least not logically impossible to think and live like a reductionist, and he points to the Buddha as an example.²⁷ While difficult, he argues, it is at least possible to strive for 'Liberation from the Self'. This may well be true, even on the extreme interpretation. It may even be true that some Buddhists have indeed managed to live according to the extreme claim. These people do not have the same intuitions and judgements as Rawls about justice; they do not have a 'conatus of desire, project and interest' of the kind by which Williams is propelled forward; and they do not deliberate or author action like Korsgaard and Blackburn do. Do we have an argument that proves that they should nonetheless accept the moderate claim? Granted, here the practice theory has no such proof available. We cannot avoid appealing to certain pre-theoretical practices, drawing out the principles

²⁶ See the discussion in S. Kagan, *The Limits of Morality* (Oxford: Clarendon Press, 1991) pp. 11-15.

²⁷ D. Parfit, *Reasons and Persons*, p. 280.

and ideas that are immanent in them, and recognising that those who do not share these practices and intuitions may forever slip through the fingers of our theories. But the claim that such a way of life – if we can call it that – is *logically possible* achieves little else. To begin with, once we realise how much depends on the moderate claim, it is plain how unbelievably far removed this possibility is from human reality. Not all people share Rawlsian intuitions about fairness, but where are the hordes that know of neither personal interest nor deliberation? Here, we should probably follow Aristotle's example and claim that such beings are either beast or God, but certainly not human.²⁸ In other words, while it is one thing to show that a way of life is logically possible, it is quite something else to claim that it is also an *attractive* possibility. How could the extremist ever convince us?

4. From Identity to Integrity

I have argued that, on the basis of the practice theory, the moderate claim is demonstrably superior to the extreme claim. However, the theory that emerges from this discussion, although in some sense rooted in the psychological principle of personal identity, should still be distinguished from the Lockean view in one important sense. Locke defined the identity relation for a person in terms of memory connections. Hume, Parfit and others are less choosy. They propose to extend the scope of the criterion from memory connections to the temporal connectedness between psychological events understood in a more general manner, including the relations between beliefs, intentions, desires, values and character traits. One idea is that two mental events are relevantly linked if they are qualitatively similar. Thus, if a desire or belief at t is reiterated at t_1 , both person-stages are connected in the way of a person. Another suggestion is that the relevant conditions are given by the causal connectedness between

²⁸ Aristotle, *The Politics*, in: S. Everson (ed.), *The Politics and the Constitution of Athens* (Cambridge: Cambridge University Press, 1996).

two psychological states. However, if we want to follow the moderate claim, we have to propose a different account of this psychological relatedness. We need a view on which the relevant identity conditions specify a *normative* relationship between the stages of a person. According to this normative account of personal identity, psychological events are internal to one person if they are reason and obligation-giving to each other. I shall refer to this relationship with the term *integrity*.

In the remainder of this chapter, I want to explain this notion of integrity in greater detail. I want to begin, however, by setting aside the issue of the identification of persons over time. For the idea that one mental state could be normative or authoritative for the other is commonly linked with the conditions of personhood or agency *per se*. Most accounts of personhood hold in common that the concept refers to a system to which intentional states can be plausibly ascribed. Agents are not subject to *external* causes, but their actions spring from a set of desires that are *internal* to the agent. However, it is much less clear whether or not every desire should be counted as internal to the intentional system. To qualify for personhood, it is often argued, intentional systems also need the capacity for self-reflection.²⁹ That is, a desire is claimed to be internal only when the agent is able to evaluate and endorse it from a higher-order level. To flesh out this idea of personhood, I briefly want to consider the following two models.

4.1. *The Hierarchical Model*

The core feature of the first model is the distinction between ‘first-order volitions’ and ‘second-order volitions’, introduced in a seminal article by Frankfurt.³⁰ Briefly, Frankfurt’s suggestion operates on the common enough assumption that humans primarily act on the basis of desires or inclinations or ‘pro-attitudes’. However, we

²⁹ See for example P. Pettit, *The Common Mind*; D. Dennett, ‘Conditions of Personhood’, in: A. Rorty (ed.), *The Identity of Persons* (Berkeley: University of California Press, 1976) pp. 175-196.

³⁰ H. G. Frankfurt, ‘Freedom of Will and the Concept of a Person’, *The Journal of Philosophy*, Volume 68 (January 14, 1971), pp. 5-20.

clearly also have the capacity to evaluate these first-order volitions. I have the desire to stop writing and to go out to enjoy what is left of fine autumn day. But I have also formed the view that it is not appropriate to translate this desire into action. I also desire to finish this paragraph. And I have judged the latter desire to be more important or commendable than the former. So I have compared and weighed up two first-order desires and I have been able to do so because I have second-order views, beliefs and perceptions regarding the desirability of first-order desires.

The decisive argument for Frankfurt is that an action can be assigned to an agent or a person, only if it issues from a higher-order volition. The idea behind this is that we often feel a distance to certain first-order desires in the sense that we do not necessarily see them as 'mine'. We have all sorts of cravings and impulses, but clearly we do not always want to act on them. There is then a sense in which my desire to take the rest of the day off is *external* to my will because I do not in fact want to act on it. A first-order desire counts as a desire of mine only if it finds backing in a second-order volition. Thus, I need to step back and embrace it by forming a judgement about the desirability of that desire. When I do so, the desire becomes *internal* in the sense that I identify myself with it.

On the hierarchical model, then, we need a second-order volition to properly authorise that we act on a particular first-order desire. It is difficult to deny that people have such second-order volitions. But several philosophers have nonetheless expressed an important reservation about Frankfurt's analysis, objecting to the authority that second-order volitions purportedly possess over first-order mental states.³¹ Why do second-order volitions count as internal? This objection is nourished by Frankfurt's original discussion, in which second-order volitions appear like ordinary, first-order desires in every other sense but the fact that they are of a higher order. They are just

³¹ Most famously, G. Watson, 'Free Agency', in: G. Watson (ed.), *Free Will* (Oxford: Oxford University Press, 1982), pp. 96-110. See also P. Pettit, *A Theory of Freedom*, Chapter 3.

preferences about preferences. However, there are no restrictions on the content of or the procedures for forming such volitions. He writes:

... a person may be capricious and irresponsible in forming his second-order volitions and give no serious consideration to what is at stake. Second-order volitions express evaluations only in the sense that they are preferences. There is no essential restriction on the kind of basis, if any, upon which they are formed.³²

Put this way, we could still distinguish second-order preferences from first-order preferences. But there seems to be no deeper ground for preferring the one to the other. Consequently, the hierarchy collapses. Frankfurt seems to presuppose that second-order volitions are always internal to the agent, but this assumption is ultimately arbitrary. His approach, then, stumbles over the objection that it fails to justify the overriding authority it grants to second-order volition. The latter desire turns out to be 'just one more wiggle in the psychic stew'.³³ Of course, we could introduce 'third-order volitions' to patch up the theory, but the main worry is that we end up in an infinite regress. In order to side step the problem that third-order volitions might also be external to the agent, we need to introduce higher-order volitions *ad infinitum*.

Frankfurt argues that we can break through the regress by claiming that higher order volitions become internal if the agent 'decisively' identifies with it.³⁴ So the appropriation of action by an agent is possible on the condition that the agent conclusively decides to embrace a first-order desire as his will. But the whiff of circularity refuses to go away. The problem is that we remain uncertain about what goes on behind the process of deciding to identify with a particular volition. We still do not know whether or not it is the agent who truly decides on this identification. In other words, we need to make sure that the decision to identify with a particular desire is itself

³² H. G. Frankfurt, 'Freedom of Will and the Concept of a Person', p. 13 (footnote).

³³ As M. E. Bratman puts in his 'Reflection, Planning, and Temporally Extended Agency', *The Philosophical Review*, Vol. 109, No. 1 (January 2000), pp. 35-61, at p. 38.

³⁴ H. G. Frankfurt, 'Freedom of Will and the Concept of a Person', p. 16. Frankfurt has refined his position further work. See his 'Identification and Externality', pp 58-68, and 'Identification and Wholeheartedness', pp. 159-176, in: H. G. Frankfurt, *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press, 1997).

internal. Thus, we would need a decision to identify with a decision to identify with a desire, and so forth.

The idea that we can distinguish between external and internal desires solely by means of a hierarchy of volition has thus been forced to retreat. The fundamental problem of regress is that

... the agent's role cannot be played by any mental states or events whose behavioural influence might come up for review in practical thought at any level. And the reason why it cannot be played by anything that might undergo the process of critical review is precisely that it must be played by whatever directs that process. The agent, in his capacity *as agent*, is that party who is always behind, and never in front of, the lens of critical reflection, no matter where in the hierarchy of motives it turns.³⁵

We can *never* solve the problem of regress by pointing to some antecedent mental event, a desire or a decision pertaining to *my* willing to identify with a certain mental event or desire.

4.2 *The Reason Model*

Velleman suggests that we modify the model more radically. His account distinguishes between external and internal desires in terms of *reasons*. To begin with, however, Velleman focusses on the mental events or desires that are 'functionally identical' to agency. As I understand him, the idea is that we isolate that mental state, the desire or aim, without which we would not be able to be agents at all. Consequently, it makes no sense to describe this desire or volition as external to the agent. Perhaps this can be explained in a slightly different manner. Velleman claims that the agent 'is that party who is always behind, and never in front of, the lens of critical reflection'. So what we need to track down are the mental states or desires that cannot be subjected to further scrutiny and that do not beg the question whether or not *I* have embraced them antecedently. The only mental states or desires that fit this description are the states or

³⁵ J. D. Velleman, 'What Happens When Someone Acts', *Mind*, Vol. 101, Issue 403 (1992), pp. 461-481, at p. 477.

desires that are *constitutive* of my agency. These desires cannot be external to the agent because it is precisely in virtue of them that there is agency to begin with.

Velleman then claims that agency must be understood as the capacity and the inclination to act on the basis of reasons. Our agency rests on the very desire to act in accordance with reasons.³⁶ Such a desire, then, never appears in front of the lens of critical reflection because it is the lens of critical reflection itself. An agent can never externalise the desire to act in accordance with reason because that would, impossibly, externalise his agency. What agency entails, in short, is a desire to commit oneself to the process of reflection and deliberation on one's desires. As agents we do not just want to act on any desire, nor just on what proves to be the strongest desire, but we want to be able to give an intelligible account of *why* we desire something. It is when we are motivated by this desire to act in accordance with a reason that agency contributes to our behaviour.

According to Velleman's model, a person authorises an action if, and only if, he acts on a desire that counts as a reason for action. Put differently, the distinction between internal and external desires is drawn in terms of desires constituting reasons and desires not constituting reasons. However, if this is our approach, we require a more detailed argument spelling out why some desires count as reasons and others do not. Thus, we need to bring in some of the restrictions on higher-order volitions that Frankfurt wanted to avoid. That is, we need to isolate the property in virtue of which these volitions are not just part of the 'psychic stew', but reason-giving. Such restrictions, I think, could be delivered in two ways.

First, we could try to impose certain objective restrictions on the content of the desires of the agent. Here we need to bring in a certain standard stipulating which desires count as valid reasons for action, independently of what a person wants or deems to be reason-giving desires. This approach, then, carries the considerable burden

³⁶ Ibid., p. 478 ff.

of having to make available an agent-independent account of such standards and restrictions.³⁷ It must stipulate which desires we reasonably ought to have, irrespective of our particular engagements and projects. Second, we could try to deal with the imposition of such restrictions subjectively. Here the agent decides himself which desires are reason-giving. The trouble with this approach is that it again raises the problem of circularity. For it remains unclear why this decision should count as his.

But there is a way by which we could steer clear of these pitfalls. The sort of restrictions that I think should invite consideration are procedural and therefore largely neutral to the content of an agent's desires. Giving a reason, I suggest, merely means giving some explanation, giving some coherent and intelligible account of why we desire one thing rather than another. On this procedural approach the most pressing concern is to rule out the volitions that Frankfurt wanted to rule in, precisely on the ground that they are capriciously formed. And if this is in some sense an objective standard, I think we should be prepared to defend it as such.

Consistent with this approach, Michael Bratman argues that a desire is reason-giving when it is 'end-setting' for practical reasoning.³⁸ What is more, end-setting desires should not be conflicting with each other. In other words, there is a coherence requirement. As he puts it himself, it is necessary that one be 'satisfied' with treating a desire as reason-giving. 'One is satisfied with such a decision when one's will is, in relevant ways, not divided: The decision to treat as reason-giving does not conflict with other standing decisions and policies about which desires to treat as reason-giving.'³⁹ A desire counts as a reason, ultimately, also because it fits in a wider framework

³⁷ See J. D. Velleman, 'The Possibility of Practical Reason', in: J. D. Velleman, *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000), pp. 170-199, mainly pp. 173-180.

³⁸ M. E. Bratman, 'Identification, Decision, and Treating as a Reason', in: *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), pp. 185-206; 'Hierarchy, Circularity, and Double Reduction', in: S. Buss and L. Overton (eds.), *Contours of Agency: Essays on Themes from Harry Frankfurt* (Cambridge MA: The MIT Press, 2002), pp. 65-85. See also L. Waddell Ekstrom, 'A Coherence Theory of Autonomy', *Philosophy and Phenomenological Research*, Volume 53, Issue 3 (1993), pp. 599-616.

³⁹ M. E. Bratman, 'Identification, Decision, and Treating as a Reason', p. 201.

explaining how all our end-setting desires are logically connected. That is, we need to understand reasons on the basis of general principles or policies. Bratman points out that ‘self-governing policies’ serve to integrate our more specific end-setting desires such that we arrive at a ranking and can resolve conflicts. But they also provide constraints on the sort of desires that merit consideration at all. That is, they rule out desires. The gist, at any rate, is that end-setting desires must be endorsed by principles and policies in order to fully count as reason-giving.⁴⁰

In order to flesh out this point a bit further, I want to bring into play the specific brand of Kantianism that Korsgaard sets out in her Tanner Lectures – though it (initially) appears, perhaps, as a somewhat bleached version of what Kant’s theory is usually thought to entail.⁴¹ On her interpretation, the idea of the Categorical Imperative captures precisely what we need. Kant framed the issue of agency or the authorship of action in terms of autonomy and the free will. But, according to Korsgaard, a will could be free in the Kantian sense only if constituted in a general form. Agency comes in the shape of applying a regulative principle to your choices. As was noted by Hegel, the perennial difficulty is that we are unlikely to be able to say much about the *content* of the maxims that should guide a person’s action.⁴² What we can say, however, is that action needs to rest on a justification that has the *form* of law. Action is authorised if validated by a principle that covers all my actions, or to put it in Kantian terms, if I act on the basis of a maxim that I could will as a general law governing *my own* actions. At least, this is what Korsgaard’s rendering of the Categorical Imperative tells us to do.

⁴⁰ On this view, what appropriates action is a certain kind of decision mechanism linked with the idea of giving a reason for action. Agency is identified with the presence of that mechanism, not with metaphysical freedom. It may be important to note that this mechanism or the principles or policies from which action must issue are not themselves required to be exempt from the causal chain. In fact, it must be true in some analytical sense that they are not entirely of our own choosing. However, once in place they nevertheless enable us to embrace or reject desires as reason-giving. And according to this theory, this explains agency. For a more elaborate defence of this view, see S. R. Wolf, *Freedom within Reason* (New York: Oxford University Press, 1990).

⁴¹ C. M. Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), in particular pp. 94-100. I should point out that I am not so much interested in the exegetical accuracy of Korsgaardian reading of Kant as in how the argument answers the question at hand.

⁴² G. F. W. Hegel, *Philosophy of Right*, Knox, T. M., (ed.) (London: Oxford University Press, 1967).

The Categorical Imperative, as she understands it, articulates what it is to be a deliberative agent. It does not specify what it is to be a moral agent. It is only when we act on general precepts that we act on reasons, and it is only when we act on reasons that we are persons.

Now, it is important to stress that the restriction thus put in place is *entirely* procedural, and therefore much less ambitious than full-blown Kantianism. However, it is not hard to see what tighter restrictions might look like. Part of the customary attractions of Kantian morality is that it is grounded in the notion of agency as such. According to this view, the Categorical Imperative takes us all the way to the moral law. In order to become agents we need to be guided, not just by general laws, but also by those policies that actually fall under the moral law. That is to say, we are persons, and thus capable of action, only if our higher-order volitions have the property that they could be willed as a universal law, not just governing *our own* action, but those of *all agents*. Under this restriction, it is an analytical truth that all persons are moral, because being moral is precisely what constitutes our personhood. The critical difference between the former interpretation and this moralised version lies in the specification of the domain over which reasons are universal.⁴³

I shall come back to Korsgaard's treatment of this issue in Section 5 below. For now, just note that on her view, agency is constituted by a set of codes, principles and laws that guide action. Like Velleman and Bratman, she primarily understands action as flowing from principles or the deliberative standpoint. However, this standpoint, she goes on to point out, also gives us our identity. It provides us with an answer to the question 'Who am I?'

When you deliberate, it is as if there were something over and above all of your desires, something which is *you*, and which *chooses* which desire to act on. This means that the principle or law by which you determine your actions is one that you regard as being expressive of *yourself*.⁴⁴

⁴³ C. M. Korsgaard, *The Sources of Normativity*, pp. 98-100.

⁴⁴ *Ibid.*, p. 100.

I want to describe this kind of coherence with the term integrity. Integrity refers to the normative order that is constitutive of the person. Without it, we would be unable to act in a unified way. As Korsgaard notes, we would not be *anyone* at all. The colloquial meaning of the term also captures another important aspect of the idea of the person here proposed. Integrity is normally thought of as a virtue.⁴⁵ When we say that someone has integrity we usually mean that he is a dependable person, someone who is true to his convictions and values, and refuses to buckle under certain pressures. Thus, consistent with the practice theory canvassed above, integrity is something that we should aim to realise. We do not always act in perfectly principled ways, but we should recognise that integrity is an important standard of conduct. We are expected to be able to justify our actions in terms of principle and coherence.⁴⁶

5. Integrity and Morality

I now return to the distinction between integrity and morality. Korsgaard argues that personal identity is a source of normativity, by which she means that it explains why we ought to do some things rather than others.⁴⁷ This claim may be correct. But it is correct in a somewhat trivial sense, as we already defined personality in terms of integrity, and integrity in terms of action-guiding principles and standards. However, Korsgaard takes a more controversial step. Ultimately, her view is that integrity also provides us with the

⁴⁵ This notion of integrity is, of course, strongly reminiscent of Plato's concept of the just soul in *The Republic*, wherein reason governs spirit and passion. Without this order conflicting passions struggle endlessly for control, thus rendering us incapable of doing anything at all.

⁴⁶ Chapter IV below proposes a view on which integrity is not understood as a kind of mind state or a thought that predates an action but emerges through the constructive interpretation of past activity. The point is that integrity should apply as a regulative ideal, a standard of conduct that persons recognise and apply to their deliberations.

⁴⁷ C. M. Korsgaard, *The Sources of Normativity*, pp. 100-104. As she puts it there (p. 101): 'Your reasons express your identity, your nature; your obligations spring from what that identity forbids.'

answer to the question ‘Why should I be moral?’⁴⁸ In fact, on her treatment of the issue, the distinction between integrity and morality seems to disappear entirely.

I want to argue that her efforts to erase this distinction are mistaken. However, begin by noting that Korsgaard does not appeal directly to the Categorical Imperative.⁴⁹ The hitch in Kant’s deduction of morality from agency, she concedes, is that it fails to explain why agency requires that we extend the domain of the Categorical Imperative to the universal level. The question ‘Why be moral’ is ultimately only answered by what she refers to as our ‘practical identities’. In other words, what we ought to do is contingent on who we are. Consequently, she cooks up an additional, identity-based argument to get Kantian morality off the hook.

Korsgaard begins by pointing out that people foster many different practical identities, mostly depending on the social circumstances they find themselves in. We are citizens belonging to a particular state, servants to a particular God, soldiers bound by codes of strength and honour, and so forth. However, over and above these ‘contingent’ and ‘local’ identities, we have to conceive of ourselves as the members of ‘the party of humankind’.⁵⁰ And she goes on to claim that this universal identity is not contingent but essential. Ultimately, we must identify ourselves as bare human beings. This is, she says, ‘simply the truth’.⁵¹ And even though we are not always fully aware of it – presumably because we are not smart enough – it is a truth that we cannot escape from.

The argument is that we must identify with a principle under which we value human agency *as such* if we are to value anything else at all. Our local identities may give us a reason for doing this or that, but having such local identities, and acting on them, only makes sense on the antecedent condition that we value our humanity as such. As she writes: ‘It is the value we place on our humanity that stands behind our other roles and

⁴⁸ C. M. Korsgaard, *The Sources of Normativity*, pp. 7-10.

⁴⁹ *Ibid.*, pp. 98-100. See Section 3 above.

⁵⁰ *Ibid.*, p. 117.

⁵¹ *Ibid.*, p. 123.

imparts normativity to them.’⁵² What is more, and this is the critical step, once we recognise our own agency as an end in itself, then we are forced on pain of logical inconsistency to appreciate everyone’s agency in a similar way. Thus, an important part of the argument is that the process of reasoning as such is essentially public, presupposing the presence of and interaction with certain other agents.⁵³ At any rate, from the fact that we must value ourselves as agents, she deduces that we must award basic value to all other agents and equally respect their reasons. It is this argument that is supposed to deliver the normativity of morality.

But I fail to see why this identity-based argument is any less debatable than a direct appeal to the Categorical Imperative. The received objection to the Kantian deduction precisely calls into doubt the validity of Korsgaard’s last step. I may have to value my agency if I am to value specific commitments, but why does this imply that I have to value human agency across the board? Why could I not apply codes, rules and standards to my life that are distinctly immoral, as Cohen suggests in his example of the Mafioso.⁵⁴ The question is still whether or not Vito (or similar ‘hard cases’), who identifies with and strictly abides by the unsavoury principles and the codes of honour of the mob, could do so with full integrity. Korsgaard needs to show that the answer is a resounding ‘No’.⁵⁵ She needs to show that Vito is mistaken about who he really is, such that, after having given the matter some further thought, he will come to accept the moral law, his true identity. But again, I fail to see any logical argument that clearly proves that a person must be moral to qualify as a person as such. And if the answer is

⁵² Ibid., p. 256.

⁵³ Ibid., pp. 131-166.

⁵⁴ G. A. Cohen, ‘Reason, Humanity, and the Moral Law’, in: *The Sources of Normativity*, pp. 167-188, at pp. 183-184. See also B. Williams, ‘Internal and External Reasons’, in: *Moral Luck: Philosophical Papers 1973-1980* (Cambridge: Cambridge University Press, 1981), pp. 101-113; *Ethics and the Limits of Philosophy* (London: Fontana Press, 1994), third impression, Chapter 4.

⁵⁵ Instead of being resounding, Korsgaard’s reply to this objection in *The Sources of Normativity*, pp. 254-258, is ambiguous. She argues that the Mafioso *should* come to see, upon deeper reflection, that this local identity harbours an inconsistency and should therefore be given up. ‘But suppose – as is likely enough! – that he never does work all this out?’ If I understand her correctly, Korsgaard admits in this case that the Mafioso has *no* reasons to act morally. Whether morality is normative or not, she writes ‘depends on how much of the light of reflection is on’.

‘Yes’, then Korsgaard merely offers the existential hope that people will reconstruct and mould their identities such that impartial morality emerges as normative.

There is another problem with this idea. The upshot of the Korsgaardian approach is that morally bad actions do not count as actions at all. Korsgaard welcomes as a virtue of her theory that

An agent cannot simply set aside the question whether his action is universalizable or just, for if he falls too far short of the internal standards for actions, what he produces will simply not be an action. In effect this means that even the most venal and shoddy agent must try to perform a good action. Performing a good action and performing an action are not different activities.⁵⁶

However, we cannot just start by ruling out actions on the basis of their moral shortcomings. Many people act in violation of the moral law – presumably – but surely we do not mean to suggest that these people are not persons and their doings not properly constituted as action? Yet this is precisely what we would have to say on Korsgaard’s theory. The implication of importing moral restrictions into this framework is that immoral and unjust action may not be action at all. We could not will evil. Whether or not this conclusion is at all plausible may be a point of some controversy. But if we choose to accept it, we are also going to have to re-examine the claim that people could be held responsible for their immoral conduct.

I want to argue that integrity is best thought of in a context that is deprived from this moral input. Korsgaard is correct, I think, to argue that integrity is constitutive of our being someone at all. But in her case this proves to be old wine in a new bottle. Loyal to the Kantian project, she still wants to show that impartial morality is the perfect fulfilment of our rational agency. In contrast, the approach I want to defend rejects this ambition. I am happy to understand integrity as a virtue separate from morality, a virtue that demands that we justify our action in terms of general rules and principles that apply to our past and future actions. While this may rule out indiscriminate and random

⁵⁶ C. M. Korsgaard, ‘Self-Construction in the Ethics of Plato and Kant’, *The Journal of Ethics* (3: 1999), pp. 1-29, p. 15.

action, there is no reason to think that integrity eliminates actions that are wrong in a stronger, moral sense.⁵⁷

6. Integrity and Time

Essentially, integrity provides us with a theory of what personhood or agency is. As agents we are expected to be able to justify our desires and actions in terms of principles that are general in the sense that they guide and cover *all* our action. In other words, we have to be able to justify our actions in terms of their rational coherence. Integrity is a virtue separate from morality, but it nevertheless forms a regulative ideal.

I now want to reintroduce a temporal element in the argument. When we conceive of persons as acting on reasons as we have above, this also facilitates a more specific and normative understanding of the relatedness between person-stages. In order to show that integrity extends through time, we need to argue that these action-guiding principles, laws and policies could also direct our actions across time. As Michael Bratman explains in his book *Intention, Plans, and Practical Reason*, we need to understand that these precepts could be future-directed.⁵⁸ In fact, we need to see that action would be impossible to understand without such future-directed plans.⁵⁹

For example, why am I, as a rational agent, engaged in writing these sentences? One answer would be that I desire to write just *these* sentences. As I move on to write the

⁵⁷ For another 'thick', but Aristotelian approach to the issue of agency, see J. Hampton, 'Selflessness and the Loss of Self', in: E. F. Paul, F. D. Miller and J. Paul (eds.), *Altruism* (Cambridge: Cambridge University Press, 1993), pp. 135-165, and in particular pp. 151-156.

⁵⁸ M. E. Bratman, *Intention, Plans, and Practical Reason* (Cambridge: Harvard University Press, 1987). See also his 'Intention and Personal Policies', *Philosophical Perspectives*, 3 (1989), pp. 443-469; 'Reflection, Planning, and Temporally Extended Agency', *The Philosophical Review* (2000): 35-61; 'Planning and Temptation', pp. 35-57; 'Toxin, Temptation, and the Stability of Intention', pp. 58-90, all reprinted in *Faces of Intention*. All page references are to this book.

⁵⁹ Notice that there is an experience-based objection to Parfit that runs parallel to this action-based concept of temporally extended persons. See D. Carr, 'Narrative and the Real World: An Argument for Continuity', in: L. P. Hinchman and S. K. Hinchman (eds.), *Memory, Identity, Community: The Idea of Narrative in the Human Sciences* (Albany: State University of New York Press, 1997), pp. 7-25 and more generally his *Time, Narrative, and History* (Bloomington: Indiana University Press, 1986).

next few sentences, I am doing this because I have formed a new and separate desire to write just *those* sentences. In this way, I have to make a separate and new choice for every single sentence, every single word, that I write today. At some later point, I will probably stop writing, I then will decide to walk to the station, and only after having arrived there, I will choose to board the train. Half an hour later, I will make up my mind to get off the train and buy some things in the local store. At home, I will cook a meal, watch some telly, and so forth.

Now, perhaps this explanation is not entirely off the mark. For obvious reasons it must be the case that every action is motivated by a present desire. An earlier desire or intention cannot govern action like a 'ghostly hand' ruling from the past, as Bratman expresses the point.⁶⁰ Having said that, the image of a sequence of temporally discrete decision-moments has little intuitive appeal. The picture that emerges here has been described as that of a 'time-slice agent', an agent who permanently starts the process of deliberation from scratch.⁶¹ The key is that for time-slice agents, the actions of his past and future selves are like those of other people. Past and future actions are like natural facts to which he can react in particular ways. They can be anticipated, but never controlled.

What is the alternative model? Bratman argues that we have to turn to a holistic mode of thinking he refers to as 'the planning theory of intention'. Essentially, his model advances a kind of reasoning that (1) plots, compares and evaluates alternative *series* of actions, each covering a sustained period of time, and (2) takes this plan or evaluation to be normative or reason-giving for each of the temporally separate selves. The first component of the model requires a particular attitude to the *future*. The second component comprises a particular attitude to the *past*.

Beginning with the future-directed component, it is surely not the case, in the above example, that I write each single sentence just for its own sake. These sentences are

⁶⁰ M. E. Bratman, *Intention, Plans, and Practical Reason*, p. 5.

⁶¹ *Ibid.*, p. 35.

meant to form an integral part of a wider argument, which is part of an even wider structure. The point is that the ‘present-directed’ intention to write this current sentence is interwoven with the broader and ‘future-directed’ intention to generate a more extensive thesis. That is, the former is subsidiary to a future-directed plan in which this sentence has a specific role and meaning. In fact, it only makes sense to engage in writing this paragraph given that I also intend to achieve a greater project.

A similar kind of reasoning is involved as I set off for the station. It seems inaccurate to suggest that I just intend to head for the station, while a separate decision to board the train awaits my future self only after having arrived there. For the second decision is already implicit in the decision to leave for the station. Although it is true that my future self will have to form an intention to actually board the train, both intentions are hierarchically interlinked. They are linked by the fact that I decided on the plan to spend a quiet evening in. To complete this plan, my future selves need to form the subsidiary intention at t_1 to walk to the station, then to board the train at t_2 , and so forth. But all these intentions are linked by the higher order plan to spend the evening in a particular way. The key is that this plan co-ordinates and guides my future activities.⁶² I do not have to fill in all the details in advance. Should I buy tomatoes or peppers? This I may well postpone to later. But while there is a level of discretion, I have also ranked sequences of actions. In any case, I do not look upon my future selves as other agents, whose actions I may predict but never control. I conceive them as actions that I am currently in charge of.

This brings us to the second, backward looking component of the planning model, namely that my future selves have to accept the earlier plan as authoritative. Notice that Bratman’s model recommends a kind of practical deliberation that is two-tiered. On the higher tier, I have to settle on a particular future-directed intention or preference, which is then put in charge of a series of future selves. As we just saw, here I need to make a

⁶² Ibid., Chapter 3.

ranking of alternative *sequences* of actions. On the lower tier, there are present-directed intentions pertaining to how the plan should be put into action at separate moments. However, once I have settled on a future-directed intention, the model only achieves something if my future selves will actually recognise the plan as reason-giving. To put it differently, if we are to reason according to the planning model of agency, then we should give *normative* priority to future-directed intentions over our present-directed aims and the considerations of our temporally discrete selves. For example, future-directed intentions should frame what further problems need addressing. My future selves have to focus on the instrumental question how to get home, rather than again assess the merits of the entire plan.

This implies that future-directed intention should be relatively secure from reconsideration and change. While no plan is in principle irrevocable, all plans require some degree of stability. Once we settle on a plan we need to persevere with it, for without steadfastness we would remain locked in time-slice mode. If I were to review and revise my plans at every instant, this would defeat the very purpose of having plans and projects at all. This is not to say that we cannot critically re-examine and reject our plans. Clearly, changing circumstances could invite us to re-assess some of our plans. But the idea that our plans are permanently up for reassessment and revision or ought to be so rationally speaking makes no sense either. According to Bratman, what we need is a habitual rule or a policy of plan-reconsideration that controls the stability or firmness of our plans.⁶³ Such rules or policies could run the gamut from the utterly rigid and stubborn on the one end to the indecisive and erratic on the other. But it is likely to be somewhere in between these extremes that the planning model should be able to find a reasonable solution, or so he argues.

Now, the idea that rational agents should commit themselves to a mode of reasoning in which earlier plans have the authority to govern present choice and action is not

⁶³ Ibid., Chapter 5.

uncontroversial. An obvious conflict arises if a person-stage prefers to deviate from the initial plan.⁶⁴ For example, while it has been my plan to finish this chapter, my current person-stage strongly prefers to stop writing and go out for a walk. Faced by this temptation, the key question is whether I should persist with the original plan or simply do what I prefer at that moment in time. On the one hand, the planning model suggests that I should persist in my efforts, for this model awards higher authority to my past intention. The time-slice model of agency, on the other hand, claims that the preference of my present self is ultimately reason-giving. It rejects the authority of the earlier plan to finish this chapter, as it considers this to be the mere preference of another agent.⁶⁵ Here, we need an argument that could demonstrate that the planning model is in some sense superior to the time-slice model of agency.

What could this argument be? The answer to this question brings us back to the practice theory of personal identity and the notion of integrity as a virtue. As we saw, this theory argues that we ought to be able to own up to our past actions. And it also points to the importance of the idea of the goodness of a life as a temporal whole. In other words, the planning model seems more attractive from the normative standpoint of the practice theory. The point is that we *should* conceive of ourselves as temporally

⁶⁴ Here, I am interested mainly in examples of temptation and myopia. In such cases we have set out to achieve a certain long-term end but at some point I have a momentary reason to abandon the plan. A famous instance of a similar problem, but with added complications, is G. S. Kavka, 'The Toxin Puzzle', *Analysis*, 42 (1983), pp. 33-36. On this, see also D. Gauthier, 'Assure and Threaten', *Ethics*, Volume 104, Issue 4 (1994), pp. 690-721; M. E. Bratman, 'Toxin, Temptation, and the Stability of Intention', in: *Faces of Intention*, pp. 58-90.

⁶⁵ Notice that this argument claims that I should *rationally* abandon my plan. So the issue does not just arise because of a temporal *lack* of rationality or weakness of will. Rather, the time-slice model of agency offers a genuine reason for me to abandon the initial plan. In other words, the question is deeper than the one discussed in accounts of temptation as a form of irrational myopia. Take for example the view defended by J. Elster in his book *Ulysses and the Sirens: Studies in Rationality and Irrationality* (Cambridge: Cambridge University Press, 1979). Elster is led to explore a variety of reasons and methods for binding or committing oneself to a future action. As in the case of Ulysses tying himself to the mast, he argues, we often have good reasons to constrain our future liberties in order to increase the probability that we continue to act rationally so as to assure certain benefits on the long run. But it is important to see that, insofar as Elster already assumes – pace the extremist – that it is rational to assure benefits on the long run, he begs the question at this deeper level. To put it differently, in arguing that human rationality involves 'global' as opposed to 'local maximisation' he precisely assumes the notion of a temporally extended self that the extremist calls into question. For a perspective similar to Elster's, see G. Ainslie, *Picoeconomics: The Strategic interaction of Successive Motivational States within the Person* (Cambridge: Cambridge University Press, 1992).

extended agents. I shall not go over this argument again. However, are there alternative strategies?

Consider briefly what Bratman claims. First, given the limitations of human cognition, he argues that rational agents will simply do better by avoiding a potentially costly process of continuous plan-reconsideration.⁶⁶ So the claim is that there are straightforward efficiency reasons to stick with the plan as the 'default-option' and to ignore the temporary preferences of certain later selves.⁶⁷ I could of course go on to re-think and re-evaluate my plan at each single moment, but it is far better, in terms of my own good, to first narrow down my options and then target my scarce deliberative resources to specific problems. Second, he argues that future-directed intention facilitates the temporal co-ordination that we need in order to successfully attain long-term ends. In other words, Bratman argues that the planning model tends to yield better prudential results than the time-slice model.

But we have to be careful to specify what we exactly mean by better results. Take the 'pragmatic' view that McClennen advocates.⁶⁸ What he calls 'resolute choice' entails that agents first decide between certain future-directed plans, that is, between whole sequences of actions, and then direct their future choice on the basis of this adopted plan. Like Bratman, he argues that once agents have settled on a plan, they should apply their deliberative energy to put the plan into action. Again, the claim is that resolute choice is superior in consequentialist terms to the step-by-step-approach of the time-slice model, mainly because it facilitates a greater degree of temporal co-ordination. However, McClennen goes on to back the more specific view that the superiority of

⁶⁶ See also J. Rawls, *A Theory of Justice*, p. 418.

⁶⁷ M. E. Bratman, *Intention, Plans, and Practical Reason*, p. 17

⁶⁸ E. F. McClennen, 'Rationality and Rules', in: P. A. Danielson (ed.) *Modelling Rationality, Morality, and Evolution* (Oxford: Oxford University Press, 1998), pp. 13-40, and more generally, *Rationality and Dynamic Choice: Foundational Explorations* (Cambridge: Cambridge University Press, 1990). For a similar though at the same time critical view, see D. Gauthier, 'Resolute Choice and Rational Deliberation: A Critique and a Defense', *Nous*, 31:1 (1997), pp. 1-25; 'Commitment and Choice: An Essay on the Rationality of Plans', in: F. Farina, F. Hahn, and S. Vannucci (eds.), *Ethics, Rationality, and Economic Behaviour* (Oxford: Clarendon Press, 1996), pp. 217-243.

resoluteness must be demonstrated on the basis of *mutual advantage*. That is to say, he thinks that resolute choice is rationally feasible only if it benefits each of the temporally discrete person-stages. He even claims that we need ‘a theory of what constitutes a fair bargain between one’s different, time-indexed selves’.⁶⁹

McClennen may be right to conclude that resoluteness sometimes yields *intrapersonally* optimal results, but the notion that the planning model should always be cashed out in *micro*-contractarian terms is unnecessarily limited. Ultimately, this approach leaves us trapped in the position of the extremist, namely that each intentional episode forms a separate agent, a view, as he puts it himself, that sees plans as ‘the means whereby the earlier and later self *co-ordinate* their choices in a manner that each judges acceptable.’ But if plans should be so acceptable in order to be rationally feasible, then resoluteness will not get us far. For it seems to me a vital ingredient in most of our plans that they aim to achieve certain ends at a later stage in our lives *in spite of* the fact that this goes at the expense of some person-stages.

Consider again the case where I am tempted to abandon my original plan to finish the chapter. Now, persisting with the plan is typically *not* going to be acceptable to my present person-stage. After all, why should I spend a single second in this dimly lit room if I could also be out there, in the sunlit streets? How does sticking with the plan make me better off at this discrete moment? It is entirely plausible that there is no answer to this question. By continuing to write, in fact, I sacrifice the interests of my present self to the interests of ‘plan-fulfilment’. The whole point about my resoluteness, and about seeing the plan through, is that it will pay off at the end, and that the temporary cost of the present self will be compensated by a later benefit.⁷⁰

A more promising variation on this argument, I think, is that the temporal co-ordination facilitated by the planning model yields better results for the agent than the

⁶⁹ E. F. McClennen, ‘Rationality and Rules’, p. 25; *Rationality and Dynamic Choice*, Chapter 12.

⁷⁰ I think that Bratman recognises this when he argues that we should follow through with our plans if we thereby avoid later regret at ‘plan’s end’. See his ‘Toxin, Temptation, and the Stability of Intention’, pp. 85-90.

time-slice model, but mainly when we look at the goodness of a life *as a whole*. The planning model is superior, but often only on the basis that doing better means doing better in the long run, that is, ‘as a temporal bundle of person-stages’. The standard we require here may include reference to the addition of temporally discrete well-being, and it will probably include much else, but my point is here that in order to reject the time-slice model we rely on the claim that this is the kind of standard that matters in our comparison between both models of agency.⁷¹ Another way of putting this is to say that we rely on the practice theory and its claim that it is the goodness of *lives* that matters fundamentally, not the discrete interests of *person-stages*.

To sum up, when placed in a temporal context, integrity demands a certain attitude towards the past and the future. We have to regard the intentions and actions of earlier and future selves as internal to our agency. So, within certain confines, we have to recognise that our past intentions and actions have a measure of authority over our present choices. Our present choices are constrained by what our past selves set out to do. What is more, integrity requires future-directed deliberation. We evaluate projects and lives with a long-term horizon, thus incorporating our future person-stages into our practical deliberations. To put it differently, integrity means that we place ourselves in a historical chain of action, which we recognise as a connecting framework that supplies us with normative reasons, responsibilities and obligations.⁷²

⁷¹ I discuss this standard in Chapter VI below.

⁷² Notice that the communitarian literature on the subject captures the same idea in terms of *narrative identity*. See A. MacIntyre, *After Virtue: A Study in Moral Theory*, Chapter 15. According to MacIntyre, we need to see our lives in terms of a ‘narrative quest’. He also dismisses the idea that history more broadly conceived consists essentially out of a sequence of temporally discrete action-moments, while the meanings or the narrative structures – the beginnings, turning points and endings – are imported retrospectively by the historian. For this sceptical view, see H. White, ‘Historical Emplotment and the Problem of Truth’, in: G. Roberts (ed.), *The History and Narrative Reader* (London: Routledge, 2001), pp. 375-389, and more generally H. White, *Meta-history: The Historical Imagination in Nineteenth-Century Europe* (Baltimore: Johns Hopkins University Press, 1973). See also David Carr, ‘Getting the Story Straight’, in: G. Roberts (ed.), *The History and Narrative Reader*, pp. 197-208, at p. 198; N. Carroll, ‘Interpretation, History and Narrative’, in: G. Roberts (ed.) *The History and Narrative Reader*, pp. 246-265.

Conclusion

Chapter I showed how associations could be compared to persons after all. Two arguments play a central role in this analogy. First, there is the reductionist position that personal identity over time is a relationship between psychological events or person-stages without these states being grounded in a deeper substratum. Second, there is the extreme claim that this relationship of identity carries no rational and moral significance. However, the problem with this analogy is that it levels down. Ethical individualism now faces a kind of predicament similar to theories of corporate personality.

While retaining the reductionist view, this chapter addressed the levelling down objection by switching to the moderate claim. According to this claim, the identity-relatedness between the temporal stages of a person is imbued with normative significance. Two stages are related in the way of a person if the one could be held responsible for the other, if the one could obligate the other or if the good of the one could form a reason to the other. Whereas the extreme claim conceives of identity as morally and rationally neutral – as social relations, too, are sometimes considered to be neutral – the moderate defines personal identity in terms of normative components or reasons.

The moderate claim rests on two further arguments. I first claimed that we have to embrace the practice theory to grasp why the moderate claim is superior to the extreme claim. Instead of grounding normative theory on a conception of personal identity, the practice theory begins by assuming that certain practices are essential to human life, at least as we conceive of it in the Western tradition. From these beginnings we then reason towards a general theory of the person and personal identity. While the extreme view may hold in a logically possible world, and while this claim may even be psychologically possible to believe, the practice theory suggests that this claim is false

in the world we live in. We think that people can own up to past actions, we believe that our future interests have special weight, and we also assume that temporal bundles of experiences matter rather than isolated experiences as such. These practices cannot be reconciled with the claim that every intentional episode should be treated as a separate agent.

I then offered a more elaborate conception of personal identity along the lines of the moderate claim and the practice theory. According to this view, the identity conditions of a person are met when each person-stage avowed the authority of general principles or intentions. Integrity lies at the heart of the very idea of personhood as such. This notion also explains how past, present and future are integrated in a conception of personal identity.

We might conclude, then, that in this way individualist theories could successfully repel the Parfitian objection. Rationality and morality do not have to become more impersonal. Having said that, the move that ethical individualism has to make to save its own neck is not without consequences. In the remainder of this thesis, I want to show that the two key arguments on which it relies, that is, the practice theory and the notion of integrity, could be used to expand the scope of personality to groups. The irony, then, is that these arguments recoup ethical individualism but facilitate a move towards corporate personality at the same time. This is I think the two-horned dilemma that ethical individualism faces. If it embraces the first horn, it exposes itself to the extreme claim and the view that each temporal person-stage is an agent in and of itself. If it chooses the second, it cannot quite so easily shrug off the idea of corporate personality.

CHAPTER III

CORPORATE PERSONALITY: A POLITICAL THEORY

If we consider my unity with myself at different times as the limiting case, we shall find it very hard to establish a difference of principle between the unity of what we call one mind and that of all the 'minds' which enter into a single social experience.

Bernard Bosanquet¹

Introduction

By invoking the practice theory and the notion of integrity it can be argued that personal identity, while a relational fact, should be conceived of as a normative relationship. That is, personal identity is defined by mutual obligations and the authority of an intention or principle over the temporal stages of a person. At the same time it becomes apparent that integrity is not something we just have or not have, but also a virtue and an ideal. We at least need to recognise integrity as a normative standard of conduct. Thus, the person may still form the fundamental unit in normative analysis, in spite of the reductionist truth about personal identity as Parfit sees it.

With these arguments in hand, I now want to return to the idea of the corporate person. The standard case against corporate personality is articulated by claims that mirror Parfit's objection to ethical individualism:

(Q1) Reductionism about groups is the familiar claim that the actions and intentions of collectives like states, nations and corporations are ontologically derivative from the actions and intentions of their individual membership. States are not grounded in some substratum, existing underneath its individual members, as some corporate realists may have mistakenly thought, but merely emerge from individual intentions and acts and the relatedness between them.

(Q2) Ethical individualism asserts that collectives, conceived of as a relationship between individuals, only have derivative significance in morality and rationality. In contrast to the temporal stages of a person, the fact that two individuals are in

¹ *The Philosophical Theory of the State*, fourth edition (London: MacMillan, 1958), p. 166.

some sense connected contains no moral and prudential significance as such. Instead, the significance and value of relational facts should ultimately be derived from their significance and value to the individual good.

Corporate personality is conceivable only if at least one of these claims could be rejected. Traditionally, perhaps, this has led us to deny *Q1*. But what I want to suggest is that we reject *Q2* in favour of its counterpart:

(*Q3*) Associativism argues that even though the actions and intentions of collectives are reducible to a relationship between individuals, it could still form a non-derivatively significant unit in normative theory. The fact that two individuals are in some sense connected could be just as significant for rationality and morality as the fact that two person-stages are so related. So, associativism argues that interpersonal relatedness is inherently reason-giving, that is, its value does not always have to be reduced to the interests and desires of separate individuals.

Why accept associativism? To answer this question I want to use a line of argument similar to the one supporting the moderate claim. So we need to show that the same pivotal arguments could be made in the case of interpersonal relations.

To begin with, again we need to invoke the practice theory. That is, we need to start by assuming that certain practices are an inescapable feature of human life or of our form of life, and from this reason towards a theory of personal identity. As I explained in Chapter II, there are several normative grounds for accepting *P3*. But now we must demonstrate that there are also normative reasons for accepting *Q3*. And individualists will probably want to reject this claim. They will want to support the moderate claim and reject associativism at the same time. Consequently, we need to provide an independent argument for the necessity of *corporate* personality over and above *individual* personality. Such an argument must illustrate that a purely individualist interpretation of the idea of the person, leaves an important gap in normative theory. I try to put together one such argument in Chapters V, VI and VII.

Second, we also need to claim that *P3* and *Q3* refer at a more fundamental level to one and the same normative kind of relationship – what I call integrity. The challenge, then, is to show that integrity is not necessarily confined to associative links between ‘person-stages’ or ‘person-events’ contained in one human life span. Integrity, it will

have to be argued, is also capable of covering a whole range of associative links between individuals, including even political association.

I want to confront the latter challenge first. In this chapter, I argue that the idea of integrity as a virtue is as applicable to interpersonal association as it is to the association between 'person-stages' or 'person-events'. But let me first provide an overall agenda for the next five chapters. The argument for corporate personality needs to remain within four critical side-constraints.

First, it should be able to explain corporate personality within the limitations of reductionist metaphysics. I call this the *reductionist* constraint. This notion should therefore be framed in reductionist terms of 'shared' and 'collective' intentions and preferences.

Second, it should nevertheless give us an account of corporate personality that is independent from the normative point of view. I call this the *sovereignty* constraint. Corporate interests, preferences and intentions, I want to argue, are not merely the summation or function of individual interests, preferences and intentions. The concept of corporate personality needs to be defined in such a way that it can add something distinctive to the normative domain. Both constraints are largely conceptual and are addressed in Section 1 and 2 below.

Third, if the idea of corporate personality is to be applied to associations like the state, it should be able to avoid romantic and nostalgic conceptions of a monolithic community. I call this the *empirical* constraint. I propose an interpretive and institutional conception of integrity that accommodates the concern that our present day world is characterised by too much complexity and a diversity of identities. I tackle this problem in Chapter IV.

Fourth, and finally, corporate personality must have normative purchase. We could refer to this as the *practicality* constraint. As my argument plainly rests on the so-called practice theory, corporate personality must be able to do normative work. It must

explain something that ethical individualism cannot adequately address. This is the subject of Chapters V, VI and VII.

With this out of the way, I begin in Section 1 below by introducing a general model of corporate personality using a reductionist conception of ‘shared’ and ‘collective’ intention. Couching corporate personality in these terms, I argue, takes care of the first constraint. However, if we follow this model, the sovereignty constraint turns out to be the greater challenge. Section 2 addresses no less than three different understandings of corporate personality in which the idea collapses back into a mode of ethical individualism: ‘conventionalism’, ‘contractualism’ and ‘voluntarism’. For each of these understandings, therefore, I advance an opposing interpretation.

1. The Reductionist Constraint

I begin with the issue of how the idea of a corporate personality could be made consistent with ontological reductionism about collectives. In order to stay within this constraint, I want to introduce the notions of ‘shared’ and ‘collective’ intention and preferences. These notions avoid anthropomorphism by arguing that the intentions of collectives emerge from a rational bundling of the intentions of a number of individuals, just as the intentions of individual persons consist out of the rational bundling of intentional episodes or person-stages. Essentially, the intentions of a group are viewed as the intentions *shared* by its members, to apply Bratman’s phrase.² Without getting trapped in too much detail, I argue that the notion of a shared intention or preference meets the first challenge.

It is important to immediately note that there is something strange about the form of the intentions that are to be shared. While occurrences in the brains of individual

² This concept of ‘shared intention’ is set out in: M. E. Bratman, ‘Shared Cooperative Action’, pp. 93-108; ‘Shared Intention’, pp. 109-129; ‘Shared Intention and Mutual Obligation’, pp. 130-141; ‘I Intend That We *J*’, pp. 142-161, reprinted in: *Faces of Intention*. All page references are to this book.

people, the key is that they need to be designed not for individual action but for group action. Another way of putting it is to say that while individuals always do the thinking and the intending, they can think and intend in different modes. More precisely, people can think in an individual and a group mode. When individuals think in the former mode, so-called '*I*-intentions' are put in charge over their decisions and actions. When they think in the latter mode, so-called '*we*-intentions' are put in charge over their decisions and actions.

The notion of a *we*-intention (or collective intention) needs to be unpacked a bit further. What does it mean for action to be guided by a collective intention? On Bratman's account of shared intention, a *we*-intention is an intention to undertake something *together*. I can intend that *I* do this or that – in this case I think in the mode of an individual – but I can also intend that *we* do this or that. In that latter case I have a collective intention. The hallmark of such intentions, then, is that they are defined over the actions of a group of people. For example, assume that John, Paul, George and Ringo intend to form a band and compose and play music together. Here, John does not just intend to play himself, he intends *them* to play *together*. And Paul, George and Ringo *we*-intend the same.

John Searle advances a similar notion of collective intention. As he puts the idea: 'We simply have to recognize that there are intentions whose form is: We intend that we perform act A; and such an intention can exist in the mind of each individual agent who is acting as part of the collective.'³ The point Searle emphasises is that such collective intentions are as primitive or as basic as so-called *I*-intentions.

Robert Sugden focuses on collective preferences rather than collective intentions. But the former have precisely the same features as the latter. According to Sugden, I can define a preference in order to guide *my* actions. But I can also define a preference aiming to guide *our* actions. If the former sort of preference directs my reasoning, I

³ J. R. Searle, 'Collective Intentions and Actions', in: P. R. Cohen, J. Morgan and M. E. Pollack (eds.), *Intentions in Communication* (Cambridge M.A.: The MIT Press, 1990), pp. 401-415, at p. 407.

think as an individual. When the latter sort of preference directs my reasoning, I ‘think as a team’. He writes:

... although team-directed reasoning is carried out by individuals, it is not an instance of individual reasoning as that is represented in the standard theory of rational choice. The two kinds of reasoning are different in structure. In the standard theory, the individual appraises alternative *actions by her* in relation to some objective (her preferences), given her beliefs about the actions that other individuals will choose. An individual who engages in team-directed reasoning appraises alternative *arrays of actions by members of the team* in relation to some objective.⁴

Now, since we engage in collective action on a daily basis, it is not implausible to think that collective intentions and preferences are everyday phenomena. However, the idea of collective intentions and preferences invites an important logical obstacle, which needs to be removed.⁵ Let us concentrate on we-intentions. Evidently, at a physical level John, Paul, George and Ringo can only act individually. There is, of course, no separately existing band that could generate music by itself. And this fact points to a discontinuity between intention and action. Collective intentions can only result in individual actions. But is it really intelligible to say that John we-intends that he *and* Paul do something? After all, we have to assume that Paul, not John, is ultimately in control of his own actions and will have to make up his own mind whether or not to participate. How, then, could John intend for Paul? And to what use?

Bratman shows in my view persuasively how the discontinuity problem could be overcome.⁶ John can we-intend, he argues, on the basis of the further belief, assumption or the prediction that Paul will have the same intention. Given that John thinks that Paul intends to do his part, he can form the intention that they undertake something together,

⁴ R. Sugden, ‘Team Preferences’, *Economics and Philosophy*, 16 (2000), pp. 175-204, at p. 193. See also R. Sugden, ‘Thinking as a Team’, in: E. F. Paul, F. D. Miller and J. Paul (eds.), *Altruism* (Cambridge: Cambridge University Press, 1993), pp. 69-89. His notion of a collective preference is similar to what Susan Hurley understands by collective agency in her book *Natural Reasons: Personality and Polity* (Oxford: Oxford University Press, 1989).

⁵ See for example J. D. Velleman, ‘How to Share an Intention’, *Philosophy and Phenomenological Research*, 57 (1997), reprinted in: *The Possibility of Practical Reason*, pp. 200-220.

⁶ M. E. Bratman, ‘I Intend That We *J*’, pp. 142-161. Velleman now accepts this response, as becomes apparent in ‘How to Share an Intention’, footnote 11, p. 204.

without assuming that he has direct discretion over Paul's actions. Whether they play together is under his control. It might be objected that the need for such beliefs and assumptions detracts from this view as compared to individual agency. But the obvious response is here that future-directed intention operates by the same logic. Similar assumptions and predictions are required in order to understand how I could logically intend in the morning to go to the cinema tonight. After all, I have currently no more discretion over my future selves than John over Paul's actions. Hence, I can only intend something in the future, on the basis of the assumption that my future selves will cooperate.

I return to this analogy between collective and future-directed intention below. There is, however, a further point that is helpful in allaying the concern about discontinuity. It is certainly true that John's we-intention has no discretion over Paul's actions. But the point is rather that John's we-intention guides and constrains his own intentions and actions in a particular way.⁷ And if Paul has the same we-intention, this will also serve to frame and constrain his own individual options in a particular way. This point is captured by Bratman's important claim that the structure of collective agency is two-tiered. A collective intention to do something *together* necessitates a degree of planning and co-ordination so that each individual member knows what he must 'I-intend' to do *himself* in order to achieve the intended result at the level of collective action. In the above example, apart from having a we-intention, John will have to I-intend that he sing a certain tune, Paul will have the I-intention that he play certain chords the guitar, and so forth. A we-intention, then, involves having a plan delegating specific tasks in such a way that subsidiary intentions are defined over individual actions and their outcomes.⁸ Collective agency, among other things, entails that lower-tier I-intentions are framed, constrained and guided by a higher-tier we-intention. To put it differently still,

⁷ M. E. Bratman, 'Shared Intention', pp. 115-116.

⁸ M. E. Bratman, 'Shared Intention', pp. 118-122; J. R. Searle, 'Collective Intentions and Actions', p. 403.

collective agency could only emerge if the collective intention or preference is ultimately in charge of an individual's decisions. Thinking as a team, an individual will ultimately have to understand the rationality of his choices and actions, in terms of this higher-order, collective intention.⁹

Three further features of this model of we-intentions need to be mentioned here. First, a collective intention or preference could still leave ample discretion for individual choice and interpretation. Rather than having to describe to the last detail what an individual member will have to do, the latter may just *constrain* our options in a more general sense.¹⁰

Second, since associations often have specific purposes, we need to demand only that individual members put we-intentions in charge over those decisions and actions that are relevant to those purposes. So group members do not altogether lose their sense of individual agency. The point is merely that we think in the mode of the group where appropriate. While we think as citizens, members of staff, and so forth, on particular occasions, in numerous other contexts we can continue to think as individual agents. What we have reason to do, then, depends on the standpoint that we take. We can either conceive of ourselves as members of a particular group or as individuals.¹¹

Third, we do not need to impose external restrictions on the content of collective intentions or preferences. In rational choice theory, individual preferences are generally taken as a given. They provide a ranking of action-outcomes, capable of guiding the agent towards one action or another. As has often been noted before, the ranking is not itself the subject of further normative evaluation. Collective preferences are similar in the sense that they provide the agent merely with instrumental reasons.¹²

⁹ J. R. Searle, 'Collective Intentions and Actions', pp. 402-406.

¹⁰ M. E. Bratman, 'Shared Intention', pp. 119-121

¹¹ For the idea that the range of the collective intention of a state should be restricted to a small number of basic tasks, see R. Dworkin, 'Liberal Community', in: S. Avineri and A. De-Shalit (eds.) *Communitarianism and Liberalism* (Oxford: Oxford University Press, 1992), pp. 205-223. See also R. Sugden, 'Team Preferences', p. 201.

¹² See R. Sugden, 'Team Preferences', pp. 197-199, and my discussion in Section 2.2 below.

Once this two-tiered structure is in place, the question is not whether John can logically intend to play while he has no discretion over the actions of Paul, George and Ringo, but whether or not they all share this collective intention and are prepared to accept its authority. It is important, at this point, to emphasise the distinction between a *shared* intention and a *collective* intention. It is perfectly possible for one person to have a collective intention. For example, John might have a we-intention to form a band with Paul, George and Ringo, while the latter do not. In this case, John believes or predicts, mistakenly, as it turns out, that they are all agreed on forming a band and that they are all agreed on a certain plan delegating more specific tasks to each of them. Until he finds out about this mistake, John will think he has a reason to proceed to carry out what the plan instructs him to do.¹³ But it is still wrong to suggest that in this example there is a group agent. An association will exist only if Paul, George and Ringo *share* this intention. And an intention is shared if two or more people both intend to undertake a certain activity together. As Bratman puts it: 'it takes at least two not only to tango but even for there to be a shared intention to tango.'¹⁴ So, the concept of shared intention refers not only to the intentions or preferences of a *single* individual, but to a state of affairs in which intentions and attitudes of *several* individuals interlock in a purposive way.

In the literature, this relation is conventionally described by a number of jointly necessary and sufficient conditions, which are the subject of ongoing debate and many subtle differences. To keep things simple, I want to introduce Bratman's account as a working hypothesis.¹⁵ Roughly, his model consists of the following conditions. We share an intention if:

(1) You and I intend that we *J*;

¹³ J. R. Searle, 'Collective Intentions and Actions', pp. 406-408.

¹⁴ M. E. Bratman, 'Shared Intention', p. 117, footnote 17.

¹⁵ *Ibid.*, pp. 116-122. For alternatives to and variations of this model, see for example R. Tuomela, 'We Will Do It: An Analysis of Group-Intentions', *Philosophy and Phenomenological Research*, Volume 51, Issue 2 (June 1991), pp. 249-277; P. R. Cohen and H. J. Levesque, 'Teamwork', *Nous*, Volume 25, Issue 4 (September 1991), pp. 487-512.

- (2) You and I intend that we *J* in accordance with sufficiently meshing plans;
- (3) I intend that we *J* partly because you and I intend so, and you intend that we *J* partly because you and I intend so;
- (4) It is common knowledge between us that the above conditions apply.¹⁶

The first condition requires that John must generally intend that they play, Paul must intend that they play, and so forth. By the same token, what Sugden calls a ‘team preference’ requires that two or more persons have the same collective preference or what he refers to as a ‘team-directed preference’.¹⁷

The second condition specifies that the parties also need sufficiently compatible plans. While it is not necessary that co-ordinating plans are filled out to the smallest detail, they should overlap at a more abstract level such that outstanding issues could be resolved in further rounds of bargaining or deliberation.

More controversially, the third condition demands a degree of respect for the individual agency of the parties involved. Using one of Bratman’s own examples, if Sonny handcuffs Donny and transports him to New York in the boot of his car, then they do not share an intention to go to New York together. Even though their actions are co-ordinated, it is not at all evident that Donny intends that they go to New York together. As Bratman puts it, this is ‘the Mafia sense’ of sharing an intention. I shall refer to this condition as the agency requirement.

The fourth condition, finally, stipulates that it is common knowledge or, weaker, a common belief between the parties that the above conditions apply. This means, as Sugden puts it, that each member must have ‘full team confidence’.¹⁸ That is to say, each individual must be confident that the other members intend to *J*, and each

¹⁶ M. E. Bratman, ‘Shared Intention’, p. 121.

¹⁷ As with collective intention, it is possible for one individual to have a *team-directed* preference. Like shared intentions, however, *team* preferences ‘are properties of a set of individuals, and require a network of common beliefs.’ See R. Sugden, ‘Team Preferences’, p. 195.

¹⁸ R. Sugden, ‘Team Preferences’, pp. 194-195. Sugden seems to subscribe to a model restricted to condition (1) and (4).

individual is confident that the other members are similarly confident, and so on indefinitely.¹⁹

Now, something like this forms the conceptual core of the idea of corporate personality. We can begin to understand how small groups and associations, and perhaps even universities, corporations and states could intend and act collectively without abandoning reductionism. To see why corporate personality is no less credible than individual personality, it may be helpful to compare the following two cases. First, let us say that you and I intend to write a philosophical study together. You and I define an intention and a preference over something we want to do together. To achieve this purpose, we need a plan for the co-ordination of our individual efforts. It is my job, say, to write the first three chapters, while it is your job to write the last three. I take care of finding a publisher; you take care of certain other things. We also need to think in the mode of a group. That is, we need to evaluate our choices and actions in terms of our joint purpose. If our subsidiary intentions interlock in a coherent manner and we have 'team-confidence', then you and I share an intention in the above sense.

Compare this to the case where I intend to write this treatise on my own. In order to succeed, it is obviously not sufficient to just form an intention to start with the first section at t . Starting the whole project also depends on my having an intention to write the next sections of the book at $t_1, t_2 \dots t_n$. What is required, then, is not just an intention to start writing now but another sort of intention covering my future actions over that

¹⁹ Interestingly, Margaret Gilbert advances a model that contains obligations and rights of commitment. See her 'Obligations of Joint Commitment', reprinted in: M. Gilbert *Sociality and Responsibility: New Essays in Plural Subject Theory* (Lanham: Rowman & Littlefield, 2000), pp. 50-70, in particular pp. 52-54, and more generally M. Gilbert, *On Social Facts* (London: Routledge, 1989). On her view, a shared intention to A emerges if two individuals 'are jointly committed to intending as a body to A '. What a joint commitment precisely includes remains somewhat unclear. What is clear, however, is that such commitments cannot be rescinded unilaterally. In contrast to Bratman's account, Gilbert's model directly provides the basis of associative entitlements and obligations. For her such entitlements and obligations are constitutive of corporate personality. I address this important issue in Section 2.3 below.

entire period of time.²⁰ So, I will have to intend at t to write further sections at t_1 and for however long I think it is going to take me to finish.

Now, the crucial point is that having a future-directed intention like this is the equivalent to having a collective intention. To begin with, it is bogged down by the same discontinuity between intention and action. I presently need to define an intention over what I will do in the future. Yet, I know that presently I have no discretion over my future actions. My intention at t has just as little discretion over what I will do at t_1 as *my* intention that *you* write Chapter 4. In other words, my intention to write a treatise myself is defined over what *we*, my future person-stages and I, are up to together.

Furthermore, and also similar to collective intention, the success of my project therefore depends on the question whether or not my future person-stages will take this intention as authoritative. At any rate, having a future-directed intention necessarily involves a co-ordinating scheme allocating specific tasks over the coming weeks, months and years. If I begin with Chapter 1 this week, I will move on to the next within two months, and so on. To put it differently, my future selves and I have to think in a style analogous to the group mode of agency. Thinking in this mode, I ultimately take my future-directed intention to be in charge over my present decisions. So, my person-stages are guided by a purpose or a preference not just defined over what I could presently do, but over the outcomes of a temporally sustained sequence of actions, over what I could achieve over a period of time. Again, it is this two-tiered model that explains my agency.

Notice also that my project will succeed, not only if my future person-stages intend to write this treatise together, but also on the condition that our subsidiary intention are sufficiently adjusted to each other. There is, of course, some discretionary space for each person-stage. And my person-stages will probably also entertain other future-directed projects. But there must be a minimum of agreement on the co-ordinating

²⁰ I discussed Bratman's planning theory of agency and the notion of future-directed intention in Section 6 of Chapter II.

scheme that distributes tasks over the coming period. Thus, my intention to write the book myself has all the hallmarks of shared intention, reductively understood. That intention is, in fact, a temporal bundle of interlocking intentions of several person-stages.²¹

Recall that I claimed that the personal identity of individuals should be understood entirely in terms of a rational bundling of purposive mind-states to which I also referred to as integrity. Integrity essentially means that we guide our actions according to a general policy. This policy must be general in the sense that it is guiding not only in the context of one concrete and temporally discrete decision, but also in a whole range of decision moments. In addition, it should be stressed that integrity is not just a kind of consistency or stability between temporal stages that, understood in terms of a system of rules or laws, allows us to predict behaviour. Integrity also involves a mode of thinking according to which this system of rules and laws is a normative standard for the appraisal of our attitudes and behaviour. So this system of rules is authoritative in the sense that it provides us with reasons for action.

Something like the above notion of shared intention precisely refers to such a bundling, the only difference being that this bundle does not consist of the temporally discrete mind-states of one individual, but of those of several individuals. As an individual but temporally extended agent, future-directed precepts like plans and intentions are put in charge of my choices and actions. In the context of corporate agency, integrity assumes the shape of collective intentions and preferences, that is, policies, plans and principles defined over the actions of a number of people and the outcomes thereof.²²

²¹ See also G. J. Postema, 'Morality in the First Person Plural', *Law and Philosophy*, 14 (1995), pp. 35-64.

²² In her *Natural Reasons* (p. 142) Hurley makes a similar point: 'The notions of personal action and a personal unit of agency already allow for the possibility and indeed the normality of the corresponding sense of intrapersonal 'collective' agency – it just is personal agency, which reflects the distinctive capacities of persons in the presence of conflicting reasons for deliberation, all-things considered judgment, and higher-order attitudes.'

2. Sovereignty Constraints

The notion of shared intention, I want to claim, takes care of the *reductionist* constraint. We do not, at least, slip back in metaphysical theories postulating a spooky mind-substance of the group. Having said that, a number of serious challenges still remain. In this section, I argue that Bratman's model – his agency condition in particular – needs to be read in a specific way to meet what I earlier referred to as the *sovereignty* constraint. Bratman's model of a shared intention is open to attack from the opposite end. That is, it may collapse back into a form of ethical individualism, thus rendering the idea of corporate personality ultimately incapable of doing any real and independent work. To prevent such a collapse, the model needs to be read and fleshed out in a particular way.

2.1 Conventionalism

To begin with, shared intention must be carefully distinguished from what could be described as conventionalism.²³ This view holds that the notion of we-intention could ultimately be broken down in standard I-intentions. The model discussed above argues that collective intentions are basic. However, a *collective* intention becomes a *shared* intention only if a we-intention is jointly held, and certain supporting attitudes are interlocking in the right way.²⁴ In contrast, conventionalism argues that the idea of a collective intention itself consists in an interlocking network of I-intentions. Thus, it

²³ The most important statement of this view is D. K. Lewis, *Convention: A Philosophical Study* (Cambridge: Harvard University Press, 1969).

²⁴ Because Bratman's notion of shared intention is partly conceived of as a network of personal intentions and beliefs, Gilbert argues that this model is conventionalist in spirit. See her discussion in 'In Search of Sociality – Recent Developments in the Philosophy of Social Phenomena', pp. 154-164, in particular p. 156, in: M. Gilbert, *Sociality and Responsibility*. She contrasts it with Searle's approach. In my view, her reading is well off the mark and confuses *collective* intention with *shared* intention. Shared intention, as I understand the model, requires a complex of we-intentions, sub-plans and beliefs. It is not an alternative to we-intention, but aims to describe the social conditions that need to be in place for collective agency to fully emerge, an issue that Searle does not address.

rejects the entire idea that we first have to think in the mode of a group. It explains collective action on the basis of the normativity of individual intentions and preferences alone.

Consider again the above example. On Bratman's model, John and Paul have a *we-intention* to undertake something *together*. In addition, their attitudes converge according to a plan that co-ordinates their individual efforts, they respect each other's agency, and so on. Instead, the conventionalist picture suggests that John I-intends to sing *alone*, that he further believes that Paul I-intends to play guitar, and that all this is common knowledge between them. What is more, Paul does in fact I-intend to play guitar, and he also believes that John intends to sing, and that this is common knowledge between them. In other words, it suggests that there is here a kind of convergence between their I-intentions and certain supporting beliefs so that they could successfully co-ordinate their efforts. As I understand it, conventionalism provides a model for social co-ordination and co-operation without having to resort to the idea of collective intentions or preferences.

Is conventionalism plausible? This is of course a key question. Fortunately, we do not have to settle it right now. Here, it is important just to emphasise the difference between both models. The idea that John and Paul share an intention remains fundamentally at odds with the conventionalist picture.²⁵ Essentially, the notion of collective intention as basic entails taking a different kind of standpoint. It implies that individuals think as a group and guide their deliberations on the basis of collective intentions, preferences and purposes. The conventionalist account of *we-intention* cannot in the end be differentiated from ethical individualism. It supports the standard view that we guide our deliberations on the basis of individual intentions, preferences and purposes.

²⁵ See J. R. Searle, 'Collective Intentions and Actions'. In fact, Searle argues that collective intentions are 'biologically innate'. See his J. R. Searle, *The Construction of Social Reality* (London: Penguin Books, 1996), pp. 23-26; 37-38.

Nothing I have said thus far actually implies that this notion of corporate personality also provides a more plausible explanation than the conventionalist model for a certain range of events. As far as I can see, the position taken here is still neutral to that question. But if we are serious about the idea of corporate personality, I suggest, we will have to commit to the view that we-intentions are primitive, even though collective agency requires an interlocking of further attitudes. Of course, it must ultimately be demonstrable that this difference has implications. And in Chapter V, VI and VII, I shall argue that this is indeed the case. But for now the point is merely that collective intentions are *sui generis*.

2.2 Contractualism

I return to the concept of a shared intention. As set out above, this concept is vulnerable to a further problem. This objection results from the interpretative discretion that the agency requirement, as set out above, appears to leave. How we understand this condition turns out to have pivotal significance, and I shall be occupied with it in a large section of this chapter.

One possible and strongly normative reading of the requirement demands that shared intentions have a contractualist basis. Essentially, this reading attempts to justify or rationalise collective intentions or preferences in terms of individual intentions or preferences. In other words, a shared intention must fulfil the normative requirement that for all the members it must be 'reasonable' to join the collective endeavour. And 'reasonable' is likely to mean here that it must pay off individually to start reasoning in the group mode. Natural candidates for these normative requirements are the Pareto principle and, more restrictively, Rawlsian or other egalitarian principles of distributive justice.

The problem with the contractual view, I want to argue, is that it rules out too much.²⁶ Of course, I see that when such conditions are not met, suspicions may be raised as to whether or not the participant's agency is sufficiently respected. But it goes too far, in my view, to rule out the possibility that people can willingly have a we-intention even if the shared activity does not pay off in terms of their own individual preferences and interests. While it may well be that such conditions set out an *external* normative parameter for co-operative action, it is not persuasive to claim that individuals could only think in the mode of a group if there were good reason to do so from the individual point of view. To put it differently, I want to suggest that collective intentions are not normatively entrenched in I-intentions. Whether or not an agent has a we-intention is in my view ultimately an empirical question, no less an empirical question than whether or not agents have a certain individual intention or preference. Thus, what we need is the presence of an intention to undertake something together. This presence is a reason for action in and of itself. We do not need a further, normative test to the effect that by the lights of some criterion it must also be *reasonable* to adopt a certain collective intention or preference.²⁷

However, let me try to make the point against the contractual view in a different way. Contractualism, I have just said, is grounded in the idea that in order to respect the agency of the parties involved, collective intentions must be reasonably acceptable. And 'reasonably acceptable' is hammered out in terms of individual intentions and preferences or rather in terms of individual gains and pay-offs. But why should this be true? Perhaps the most promising argument trades on a perceived distinction between shared and collective intentions on the one hand and I-intentions on the other. The idea, it seems, is that the former intentions do not count as a form of agency, while the latter

²⁶ See R. Sugden, 'Team Preferences', pp. 186-190.

²⁷ Cf. R. Sugden, 'Team Preferences', pp. 197-199; M. Gilbert, 'Collective, Preferences, Obligations, and Rational Choice', *Economics and Philosophy*, 17 (2001), pp. 109-119, at pp. 113-114.

do, because we-intentions require us to think in a *modus operandi* fundamentally estranged from ourselves.

I believe this argument to be flawed. Take again the case where I intend to write a philosophical study on my own. As we saw, I do not just form an intention to start with the first few words. Instead, I also intend that my future selves write later paragraphs and chapters. Thus, my intention is defined, not over immediate action, but over temporally sustained action. Notice also that my future selves have to think in a mode of thought in which this future-directed intention is taken to be reason-giving. This explains why we think of our individual lives as continuous wholes rather than unconnected sums of infinitesimally brief moments. To put it differently, we have preferences and goals pertaining to how our whole lives go or at least large chunks of it. And in considering what we ought to do, we abstract from our present interests in order to also look at the good of our life as a whole.

The problem, then, is that we could also apply the contractualist thesis to individual personality.²⁸ Contractualism would demand that our future-directed intentions be reasonably acceptable from the point of view of each person-stage. My future selves would all have to be able to share in the gain or receive their fair share under some criterion of justice. But this is absurd. We do not subject ourselves to such a restricting regime. If I presently decide against going to the cinema, while I would desperately like to, and, miserably, continue to work with the ulterior purpose of completing the task at hand, there is no one who would seriously doubt my free agency. We do not say that this could not reasonably be an intention or an end of mine because my present self is worse off than he would have been without this intention. On the contrary, we often take this to be an exemplary case of rational agency that needs to be singled out for praise. There is, of course, no insinuation of unfairness either. I just happen to have this

²⁸ E. F. McClennen sometimes appears to support this view in 'Rationality and Rules' and, more generally, *Rationality and Dynamic Choice*.

long-term goal, and if I endorse that goal, that is pretty much the final word. That goal provides me with a reason not to go to the cinema.

It is important to note that I am not arguing against some distributive policy according to which, say, future-directed action must not disturb a certain allocation of well-being over my entire life-span *as a measure of the good life*. I may well have good reasons to prefer to live the egalitarian life, a life without having to suffer such miserable moments, but perhaps also a life without having written a book. Distributive considerations of this kind, I suspect, do play a role in many of our deliberations about the good life. We set limits to how much sacrifice we presently want to endure in order to realise long-term goals, even though these limits seem a good deal less restrictive than Rawls's difference principle or even the Pareto principle.²⁹ But here my argument is not with these principles. It may not be absurd to impose distributive limits on how we divide well-being over time. What is absurd, however, is the suggestion that these limits are normatively powerful enough to rule out the conceptual possibility of my *reasonably* and *willingly* adopting a future-directed intention that requires a present sacrifice for a future good. This suggestion is absurd because such normative principles are external to agency. Whatever use they have, they are not a prerequisite for agency to emerge.

In resisting the contractual reduction of we-intentions to I-intentions, I am pressing the exact same claim. I am not suggesting that there are no *external* grounds for adopting some distributive criterion such that states or other co-operative schemes can be ethically appraised. Perhaps it is true in a particular context that some egalitarian distribution is what justice demands. I have, at any rate, no real quarrel with such a norm here. But I do have a problem with its specifically contractual grounding in

²⁹ In Chapter VI, I endorse Velleman's views in 'Well-being and Time', reprinted in: J. D. Velleman, *The Possibility of Practical Reason*, pp. 170-199. According to Velleman, the good of a life cannot be entirely reduced to a simple addition of synchronic well-being. Over and above this addition, an independent distributive criterion is at work. This criterion, however, is not adequately captured by egalitarianism. Rather it suggests that the good of a life is partly shaped by its historical progression over time, by its narrative structure.

agency. It is simply implausible to claim that it is impossible for someone to *reasonably* and *willingly* agree to participate in a joint venture that violates this criterion and, say, leaves him worse off than before. Contractualism effectively argues that such a joint venture cannot be grounded in a person's will and, therefore, that there cannot be a shared intention in these cases. Now, it may not be mistaken to embed shared intentions in an individual's volition and agency. What is mistaken, however, is the presumption that only individual intentions could count as constitutive of people's volition and agency. This presumption takes the distributive requirement for an internal requirement constitutive of agency itself.

On the account I am suggesting here, we-intentions and collective preferences are no less expressive of a person's will, no less normative than I-intentions. They may directly provide us, that is, with reasons to choose and act. If John and Paul come to hold the collective preference that they play music together, then they rationally ought to do so, even if this means that they are worse off individually as the result. In this respect, I can really see no reason why we ought to treat we-intention or shared intention differently from future-directed I-intention. We do not generally think that future-directed I-intentions, like the one in my example above, deflect us from our agency simply because they require us to think in a particular mode of thought. We do not think that this kind of intention must be justifiable from the perspective of present-directed volition in order to count as a true expression of our agency. In fact, when our future-directed intentions are governed by the present-directed viewpoint we are vulnerable to what we commonly describe as weakness of will. Why should we think differently in the case of we-intentions?

I should add an important caveat. The very notion of the *weakness* of will clearly implies that our temporally discrete desires, far from being expressive of our free agency, actually threaten to undermine it. Our individual agency, according to this view, is always located in future-directed intention. In other words, by overriding the desire of

a person-stage, we fulfil rather than impede our freedom.³⁰ That said, we should be careful not to press the analogy with collective intention to its ultimate conclusion. If we were to claim that collective intention is always more expressive of our agency than individual intention, then we would end up in the position that Rousseau may have set out in *The Social Contract*, and which has provoked the wrath of classical liberalism ever since, namely that elements resisting the General Will could be ‘forced to be free’.³¹ Instead, we should emphasise the pluralist consequences of the argument. The main point I am trying to make is not that collective intention is ‘more real’ than individual intention; it is merely that it is *conceivable* that a collective intention is expressive of our agency. Individualists assume that future-directed intentions are somehow privileged, that it is somehow self-evidently true that these intentions correspond to our agency, and that our temporal selves can thus indeed be forced to be free. But there is in my view no ironclad argument to back up this assumption. As Parfit demonstrates, there is no logical reason why we could not argue for the autonomy or the rights of temporal selves, and deplore the repression that ethical individualism necessarily involves. More to the point, there is no logical reason to presuppose at the outset that collective intentions could never be truly expressive of the will, and should always be subjected to the contractualist test. What this leaves us with is not the Rousseauian position, which may indeed lead to the Gulag, but rather the view that the unit of human agency is more flexible than individualists normally assume, and could in principle take on several forms. We do not rub out the practical importance of the individual and conflicts between individual interests in favour of the interests of a collective person. But what we have to affirm is that individual and corporate personality may *both* have a role to play in practical reason. Moreover, the question

³⁰ That would appear to be a view endorsed by J. Elster in his *Ulysses and the Sirens*. He typically views the problem in terms of a *breakdown* of rational agency.

³¹ J. J. Rousseau, *The Social Contract*, Book I, Chapter 7, in: G. D. H. Cole (ed.) *The Social Contract and Discourses*.

which particular form our agency should assume is in itself the subject of normative inquiry.³²

To sum up, we must be careful, at this stage, not to place conditions on corporate agency that prove to be too strict and are, at any rate, not properly applied to individual agency either. Certainly, some conditions of agency need to be met if intentions are to be really 'shared'. But contractualism goes wrong in fixating on a norm that insofar as I can see is external to corporate personality.

2.3 Voluntarism and Obligations of Shared Intention

While I reject the contractualist interpretation of shared intention, I have not rejected the agency requirement as such, namely that such intentions must respect the volition and agency of the individuals involved. I merely want to keep it separate from distributive norms testing the *reasonableness* of collective intentions and preferences.

If the contractualist view is rejected, it might be thought alternatively that shared intention requires something like an explicit agreement, a promise or an actual contract between the parties involved. This thought is appealing because an actual contract would seem to satisfy all the conditions of shared intention set out above. The problem, however, with this view is that it involves the danger of excluding associations that do not obviously rest on such an agreement, yet should still be counted as corporate persons. The state is an obvious case in point. It is widely agreed that state membership does not rest on explicit consent or an actual contract.³³ Hence, if we want to rule in non-voluntary associations like states, the actual contract view must be dismissed as too restrictive.

³² If I understand her correctly, Hurley makes a similar point when she sets aside the erroneous view that 'the unit of agency must be fixed, somehow, once and for all'. According to Hurley, '... we must allow ourselves to survey the units of agency that are possible in the circumstances at hand and to ask *what the unit of agency, among those possible, should be.*' See her *Natural Reasons: Personality and Polity* (Oxford: Oxford University Press, 1989), pp. 145-146, and Chapter 8 more generally.

³³ For an early statement, see D. Hume, 'Of the Original Contract', in *Essays: Moral, Political and Literary*, E. F. Miller (ed.) (Indianapolis: Liberty Fund, 1985), pp. 265-487.

I suggest the following argument. For a shared intention we need two or more persons possessing a collective intention, sufficiently meshing sub-plans, and so forth, as specified above. Now, express consent or a contract or a promise are merely manifestations of the fact that such intentions are present. To put it differently, agreement is not itself part of the *criterion* for shared intentions, but merely one *indicator* among a host of possible indicators revealing the prevalence of the we-attitudes in question. If John and Paul come to an agreement, sign the relevant papers, and so forth, this may be ample to satisfy the conditions of sharing an intention. However, while it is one thing to say that the presence of an explicit agreement is *sufficient* evidence for the presence of a shared intention, it is something else to claim that this is also a *necessary* condition for shared intention to emerge.

The latter half of this claim is false because it is possible for people to share an intention without explicit agreement or without making a promise. Bratman approvingly refers to the Hume's conventionalist conception of justice.

I observe, that it will be for my interest to leave another in the possession of his goods, *provided* he will act in the same manner with regard to me. He is sensible of a like interest in the regulation of his conduct. When this common sense of interest is mutually express'd, and is known to both, it produces a suitable resolution and behaviour. And this may properly enough be call'd a convention or agreement betwixt us, tho' without the interposition of a promise; since the actions of each of us have a reference to those of the other, and are perform'd upon the supposition, that something is to be perform'd on the other part. Two men, who pull the oars of a boat, do it by an agreement or convention, tho' they have never given promises to each other.³⁴

Hume's rowers have the collective intention to row the boat together. They agree on where they are going, at what speed, and so on. In other words, they have appropriately interlocking attitudes. They are in fact prepared to row the boat. It might be claimed that their express consent is required as a means to signal to each other that they are prepared to row the boat. To be sure, consent and promises may help to facilitate mutual knowledge. But, again, it is a mistake to claim that they form a necessary means. What

³⁴ D. Hume, *A Treatise on Human Nature*, p. 490. See M. E. Bratman, 'Shared Intention', p. 111.

is necessary is that both men are confident that the other has the appropriate collective intention and supporting attitudes. However, this confidence may also rest on a different kind of evidence. As both men are in fact rowing the boat, their collective preference is revealed by their continued action. The rowers may simply base their expectations on stable patterns of behaviour. Shared intentions, then, may be different from conventions in that they rely on collective intentions or preferences, but it is nevertheless true that they could emerge in equally spontaneous and informal ways.³⁵

The express consent interpretation, in short, is unnecessarily stringent. What is more important, it does not help to alleviate the concern with the agency of the parties involved. Essentially, the problem is that shared intentions could be elicited under duress. Recall the 'Mafia sense of going to New York together'. Here, it is important to distinguish two versions of this general case. Let 'Case A' refer to the scenario where Sonny wrestles Donny into the boot of his car and drives to New York. Clearly, Donny is brought to New York entirely against his will. Since the latter does not intend to go to New York at all, it is evident that there cannot be a shared intention either.

However, consider Case B. This time Sonny does not employ his brute strength, but threatens Donny with a gun. So Donny faces a choice. It may be a harsh choice, perhaps, but he can still decide to accompany Sonny to New York or get shot in Miami. Now, if he decides to join Sonny, it may be plausibly claimed that Donny does indeed form the collective intention that they go to New York together, their intentions may also be sufficiently meshing, and it is not unlikely that this is common knowledge.

The key question, however, is whether or not Donny's agency is respected such that this collective intention could be *shared*. According to one view, it matters that Donny

³⁵ Similarly, Gilbert argues that we could come to share an intention, not by promises, but merely by persistently using the pronoun 'we' in a certain context. Thus, when we speak of 'our country', 'our laws' or 'our government', this indicates that an intention is shared at the level of the state. She also points out that many people have a sense of political obligation. If I understand her correctly, these are forms of identification that are sufficient to constitute us as a 'plural subject'. See her 'Reconsidering the 'Actual Contract' Theory of Political Obligation', reprinted in: M. Gilbert, *Sociality and Responsibility*, pp. 97-122.

has a choice, makes up his mind and intends that they go to New York together. Albeit under duress, there is a shared intention. Call this the *harsh view*. According to another view, Donny's agency has not been respected. His collective intention is declared void precisely because it has been obtained under duress. Call this the *no-real-choice view*.

Notice first that the demand for express consent or a promise does not coincide with this latter view. After all, Sonny might even make Donny sign a contractual statement to the effect that he obligates himself to go to New York. In order to share an intention on the no-real-choice view we need *voluntary* consent. Whether consent is explicit or implicit is less important. What matters is that collective intentions are arrived at without any threats.

The real question, then, is whether or not we have to subscribe to this interpretation. In providing an answer, much seems to depend on the further question whether or not shared intentions involve mutual obligations (and entitlements). Bratman and Sugden propose a model on which shared intentions and preferences do not necessarily obligate the parties involved. The fact that you and I share an intention *may* have normative consequences.³⁶ For example, it may create mutual expectations that we are obligated to meet. However, these obligations are not internal to the idea of a shared intention itself. You and I could share an intention without acquiring any obligation whatsoever. If this claim were correct, then the harsh view would be a good deal less harsh. As I understand him, Bratman and Sugden do indeed hold that in Case B there is a shared intention. But this result does not imply that Sonny has an obligation to do his share, at least not in virtue of sharing an intention.

Gilbert, on the other hand, advances a model on which such mutual obligations and entitlements are conceptually constitutive of sharing an intention – what she prefers to call a 'joint commitment' or 'joint decision'.³⁷ This idea is to operate in the spirit of the

³⁶ See M. E. Bratman, 'Shared Intention and Mutual Obligation', p. 140. Cf. R. Sugden, 'Team Preferences', pp. 189-190.

³⁷ This view appears throughout her work on the subject, but emerges perhaps most clearly in 'Obligations of Joint Commitment'.

obligations of actual contracts and promises. Yet she also supports the harsh view. According to Gilbert, if Donny decides to go to New York in Case B, this becomes a 'joint decision' of theirs. Typically joint decisions cannot be rescinded unilaterally. Thus if Sonny and Donny share an intention, this also generates an obligation on Donny's part to see their joint commitment through.³⁸

Now, Gilbert's view that obligations are in part constitutive of shared intention seems to me correct. As I argued in Chapter II, the personal identity of individual agents over time should be defined in terms of a normative and obligation-creating connectedness or integrity.³⁹ Since I now invoke the same notion of integrity in order to advance the possibility of corporate personality, shared intention has to be given a similarly normative foundation. However, intuitively it seems that something goes horribly wrong on this view. Surely, we would not seriously want to contend that Donny has a genuine obligation to accompany Sonny to New York. Either they do not share an intention at all or, if they do, they do not have Gilbert's obligations and entitlements. Thus, we are faced with a dilemma. If we want to retain the view that obligations are constitutive of shared intention, there seems to be a strong case for the no-real-choice view. If we want to maintain the harsh view, then we have to drop the inherent obligations and entitlements. Something has to give.

Gilbert suggests an interesting way to slip between the horns of this dilemma. On her account, Case B contains both a shared intention as well as mutual obligations. Her solution is to distinguish between having an obligation and having a *morally binding* obligation. The key assumption of her argument is that practical reason is pluralist in nature. On this pluralist view, Donny's obligation constitutes just one reason for action.

³⁸ See M. Gilbert 'Agreements, Coercion, and Obligation', *Ethics*, Volume 103, Issue 4 (July 1993), pp. 679-706; 'Reconsidering the 'Actual Contract' Theory of Political Obligation', reprinted in: *Sociality and Responsibility*, pp. 97-122; 'Collective Preferences, Obligations, and Rational Choice', pp. 109-119.

³⁹ Notice that in contrast to his model of shared intention, Bratman's planning model of intention is less timid in embracing this normative dimension as constitutive of our individual agency. It seems to me an essential part of his idea of future-directed intention that our time-sliced selves are obligated by this intention. You are a temporally extended agent just *because* your person-stages accept the normativity of future-directed intention. See my discussion of this model in Chapter II, Section 6.

Apart from his obligation of joint commitment to Sonny, he may have obligations of fidelity or reciprocity to other people. He may have certain general obligations of justice or benevolence. And apart from other-regarding concerns, he may also have prudential reasons. Gilbert's point, I think, is that it is perfectly possible to say that Donny is under an obligation to go to New York with Sonny, and yet to deny that Donny indeed ought to go to New York *all things considered*. Whether he ought to go in the final analysis depends on the relative weight of this obligation vis-a-vis other reasons. To put it differently, Donny has a genuine obligation, but we could claim that this obligation is trumped by more significant reasons.

It sometimes looks as if Gilbert regards the relative weight or the strength of obligations of joint commitment as a constant factor. Whether or not they are outweighed by other reasons does not appear to depend on the nature of the circumstances under which joint commitments are made. But I think it would be more accurate to suggest that this relative weight is a function, in part, of precisely these circumstances. Viewing the relative weights as such a function, something like the following picture results. The analysis of Case A remains unchanged. There clearly is no shared intention – Donny's agency is simply not effective – and consequently there are no obligations either. In Case B, Donny has a genuine obligation to Sonny in virtue of the fact that they share an intention. However, the relative weight of his obligation is significantly reduced. There is a genuine obligation, even in this case, but the weight that this obligation can put in the balance in comparison with competing obligations is relatively weak.

When I say relatively weak, I mean weaker than in Case C. In this case Donny proposes to go to New York with Sonny out of his own accord. It is less likely, on the balance of things, that Donny ought to go to New York in Case B than in Case C, or so I argue. By the same token, it is also less likely, on the balance of things, that Sonny has reason to feel aggrieved if Donny ends up not going to New York in Case B than in

Case C. To be sure, even in the latter case Donny may have stronger reasons to stay in Miami, say, because of a previous commitment to his pregnant wife. But it seems altogether more plausible that Sonny can rightfully censure Donny for breaking his commitment than in Case B. If Donny goes back on his commitment to Sonny he needs a stronger reason to trump his obligation to Sonny.

In this way we can do justice to the central intuition that something goes wrong in this view of associative obligations. Our intuitions, of course, are that such obligations and entitlements must at least be sensitive to coercion, pressure and threats. If I am right about Gilbert's view, her solution is not really responsive to these circumstances. She merely argues that such obligations *could* be trumped by other reasons, all things taken into consideration. But presumably there can be such trumping even in cases where there is no duress. No amount of trumping does justice to the central intuition, unless we can also show that trumping is in fact more likely to occur in cases where there is duress than in cases where there is not. In contrast to this, I argue for the view that while threats may not snuff out our choice completely, they do constitute a more or less serious impairment to our agency. On this approach, then, there is scope for the view that obligations are inherent in shared intention only insofar as our agency remains intact.

Alternatively, the no-real-choice view aims to capture the same intuition by arguing that shared intentions, like agreements and promises, cannot be extracted under duress. If some credible threat is issued, or so the view goes, the shared intention is void and, consequently, there is no scope for obligations of joint commitment either. What commends my approach over this view? One reason, I think, is that it does not seem very plausible to argue that joint commitment, and therefore their obligations, become entirely void as soon as a threat is issued or some degree of pressure exacted. Clearly, agreements and joint commitments are reached under the sway of threats and pressures all the time. Furthermore, what we have to bear in mind is that there are many kinds of

threats and pressures, differing in degree and intensity. For example, instead of reaching for his gun, Sonny may threaten to blow the whistle on Donny's illegal nightclub, he may threaten to pull out of some potentially lucrative business deal, in fact, he may just tell Donny he will no longer be welcome at his next birthday party. There is, I want to suggest, a fairly elastic difference between the issuing of threats on the one hand and what we would normally refer to as 'bargaining power' or 'cajoling' on the other. It would seem that if some shared intention were reached under the latter conditions, it is not entirely vacuous to suggest that these are genuine commitments that also generate genuine obligations.

The no-real-choice view could try to remedy this by capturing the complexity involved in what it means by real choice. But it would still need to propose a problematic cut-off point between joint commitment reached under the sway of legitimate cajoling and joint commitment reached under illicit duress. This refined view would declare a joint commitment void once agency is not *sufficiently* respected. At that point, concomitant obligations and entitlements simply cease to exist. There may or may not be such a cut-off point, but the point is that it is difficult to see how there could be degrees in the 'sharedness' of an intention. We either share an intention or not. Similarly, on the no-real-choice view either these obligations exist in their full glory or they do not exist at all. In contrast, my suggestion is that shared intention persists in the face of any form of pressure (with the exception of Case A). But the restriction placed on agency by this coercion is discounted in the relative weight of the resultant obligations.

That said, the difference between the no-real-choice view and the position that I just set out may seem minor. Ultimately, both views make obligations of shared intention sensitive to coercion, pressure and threats. Thus, we cannot claim that this argument answers any outstanding issues, for example, the question of political obligation, with any finality. Nevertheless, it still produces more than a merely semantic result. Two

points stick out. To begin with, if we take the no-real-choice view, then we have to accept that non-voluntary associations like states may not qualify as corporate persons, implying that corresponding obligations of shared intention could not exist either. However, there is a difference between saying that associative obligations could not exist and saying that they may be outweighed by other reasons.

Second, and more important to my purposes here, this way of approaching the issue allows us to enjoy two benefits. On the one hand, we can treat the issue as to whether a corporate person *exists* separately from the question whether or not the interlocking of attitudes has emerged under duress. This is the key advantage of the harsh view that Bratman defends. On the other hand, it still allows us to advance Gilbert's view that obligations are constitutive of shared intention, without necessarily embracing counter-intuitive results. To put it differently, we can still claim that shared intention is a normative relationship comparable to personal integrity. Of course, once the existence issue has been settled, we must still decide on the weight of associative obligations of shared intention. The presence or absence of duress is clearly one variable that has to be taken in consideration. And in terms of political obligation, this variable may have a rather limited justificatory power. However, in Chapter VII, I shall discuss other variables that could give associative obligation a more substantial weight. With this hard question out of the way, I want to conclude this chapter.

Conclusion

I argued in Chapter I and II that the identity of persons over time should be understood entirely in terms of a bundling of mind-states without reference to a physical or spiritual substratum, that this bundling should be in part defined in normative terms of integrity, and that this conception of personal integrity is itself irreducibly practical. If these arguments are correct, then we face the following two challenges to get the idea of a

corporate person up and running. First, we have to present an understanding of human association that corresponds to the idea of integrity. Second, because the model depends on its normative impact, we also need to show that this understanding of human association has practical benefits.

In this chapter, I tackled the first challenge. I wanted to show that the notion of integrity could be applied, not just to the temporal stages of a life, but also to the relatedness between individuals. To this purpose, I introduced the model of collective and shared intention. Essentially, this model explains how we could say that an association could have intentions without resorting to implausible versions of corporate realism, which postulate a group being separately from its membership. The model of shared intention relies on the view of intention as an interlocking set of collective intentions and preferences, and other supporting attitudes.

In order to further prevent shared intention from collapsing into a mode of ethical individualism, I rejected a number of interpretations. Contrary to conventionalism, corporate personality argues that collective intentions are primitive. That is, collective intentions are not to be broken down and analysed in terms of individual intentions. Against contractualism, corporate personality holds that collective intentions are not subjected to a contractualist test. Collective intentions do not have to be reasonably acceptable to an association's members. Finally, in opposition to voluntarism, corporate personality maintains that shared intentions could exist under duress, and that obligations and other normative concerns are constitutive of shared intention.

Conventionalism, contractualism and voluntarism have to be resisted because the idea of corporate personality should ultimately be capable of doing independent normative work. But before confronting this practical challenge, I want to explore the empirical dimension of corporate personality. One obvious difficulty with the above model of shared intention is that it does not always correspond well to the realities of

human association. To address this concern, Chapter IV introduces further adjustments to the model to make it applicable to the state.

CHAPTER IV

THE PERSON OF THE STATE

Introduction

In Chapter III, I offered an account of corporate personality that fits within two largely conceptual side-constraints. First, I wanted to give an account of corporate personality without having to depart from reductionist metaphysics. I called this the *reductionist* constraint. The account that meets this requirement, I argued, is framed in terms of shared and collective intentions and preferences. The corporate thesis, thus understood, is merely the extension of the notion of integrity, set out in Chapter II, to the larger scale of interpersonal relationships. Second, I also emphasised that this account of corporate personality has to be able to provide a genuine alternative to ethical individualism. I referred to this as the *sovereignty* constraint. And so I rejected conventionalist, contractualist and voluntarist interpretations of shared intention.

In this chapter, I want to introduce a third constraint, which is empirical in nature. Corporate personality, I argue, should not merely be conceivable as a logical construct; it should also be possible to claim that this idea of corporate personality applies to human associations in *this* world. Moreover, while the literature on collective and shared intention is built largely around two or three people examples, I want to argue that the idea of corporate personality could be extended to apply to human associations on a larger and more complex scale. To be more precise, I want this notion to apply to the state.

The chapter proceeds as follows. Section 1 begins by outlining the challenge. Essentially, the difficulty with extending the notion of corporate personality to the state is that this conjectures a collective intention or identity that is shared at a societal scale and neatly coincides with certain territorial boundaries. This image, now, is alleged to

be inconsistent with the complexity of associative life and the diversity of our practical identities. Briefly, the objection is that there simply is no shared intention that meets these conditions. In response, I attempt to amend the notion of shared intention in a suitable manner. Section 2 claims that we should think of intention in an *interpretive* rather than a *psychological* manner. But while successful in some parts, this approach falls short in certain others, or so I argue. Section 3 discusses two obvious but I think ultimately unpromising ways to make headway. Finally, Section 4 imports an institutional dimension into the model. It draws a distinction between *deliberative* integrity and *institutional* integrity and argues that only with the latter we obtain a better handle on the problem at hand.

1. Problems of Scale

I want to start by mapping out the challenge more precisely. *Ex hypothesi*, the notion of the state as a corporate person seems to entail at least three different things. First, it implies that its citizens have sufficiently overlapping collective intentions at the societal level. And that level usually covers a wide-ranging number of people and many highly intricate and complex forms of co-ordination, planning and so on. Second, states are usually fairly sharply demarcated by territorial boundaries. In other words, a collective intention must be shared by a precisely defined population or membership. Third, within that territory states claim a kind of authority that is alien to most other associations. While most associations could logically co-exist with other, like-minded associations, states typically cannot, not on the same territory.

Since each of these things presents a *prima facie* problem to the idea of the personality of the state, the nature of the challenge is threefold. The first problem, highlighted in an article by Kutz, arises by moving from simple to more complex

examples of collective action.¹ I shall refer to it as the problem of complexity. On Bratman's approach, individuals form a we-intention that they undertake some activity together. What this demands is an 'executive perspective', a point of view that comprehensively assesses what the group does or should do *as a whole*. On the second tier, people have to comply with a 'sub-plan', which delegates 'subsidiary intentions' pertaining to what each member should do *as an individual* by way of contribution to achieving the collective outcome. Now, the literature on shared and collective intention abounds in relatively simple cases such as: going for a walk, painting a house, and playing football. In such examples, it may well be conceivable that each of the participants first assumes this executive perspective and then constrains his actions accordingly. These cases involve a small number of people and take place in informal and discursive contexts that allow face-to-face contact, negotiation, and so on. However, on the face of it, the idea that all participants occupy the executive standpoint seems far less probable in the context of large-scale and complex associations like states. It seems downright naïve to believe that ordinary citizens or middle-ranked bureaucrats or even politicians would generally bother to consider the question as to what they are aiming to achieve collectively. Generally speaking, in such contexts our deliberations remain restricted to the second, subsidiary tier. While our actions are perhaps configured by a larger plan, it is not true that this plan is a conscious intention of each and every participant.

The second problem relates to boundaries and membership. As I understand the issue, in the literature on collective and shared intention membership consists simply in having the right intentions and meshing sub-plans. So whether someone is an insider or an outsider to a group depends on the intention and attitudes of the persons involved. John forms a group with Paul, George and Ringo because they all share an intention to play a song together. Mick and Keith remain outsiders because they have different

¹ C. Kutz, 'Acting Together', *Philosophy and Phenomenological Research*, Vol., LXI, No. 1, July 2000, pp. 1-31.

plans. The latter could become members, however, just by assuming the right group attitudes. Again, this may seem plausible in some contexts, but the membership of states differs significantly. As with many other organisations, we cannot just become members of states by forming the proper attitudes.² An additional difficulty is that states are territorially bounded. Both facts suggest that other criteria are at work.

Finally, the third challenge is formed by what John Rawls calls the fact of (reasonable) pluralism.³ This fact appeals to a by now rather commonplace image of modern life in Western societies as characterised by an increasing plurality and fluidity of practical identities, purposes, ideals and values. This diversity of ends or ‘comprehensive doctrines’ is naturally reflected into a variety and fluidity of associations and groups. Generally, it is not longer the case that people live their entire lives within the fixed parameters of one monolithic, authoritative and all-embracing community. This is not to say that modern man is solely committed to himself. Rather, individuals develop cross-cutting allegiances to many associations, to self and nation, to church and trade-union, to family, football clubs, sub-cultures, and so forth. On the Humean theory defended here, even individuals count as associations. All these associations and communities appeal to our loyalty and they all reflect particular standpoints and shared intentions that lie at the base of our practical identities. Yet it would seem that the notion of the personality of the state demands *one* project or *one* ethos or *one* identity. Thus, the problem is that corporate personality seems to depend on the obsolete picture of the state as *Gemeinschaft*, to use a concept coined by Tönnies.⁴ To put it differently, the corporate state would have to be based on the potentially oppressive mode of human association to which Oakeshott referred as *universitas*. In Roman legal discourse, *universitas* referred to ‘persons associated in a manner such as to constitute them a natural person; a partnership of persons which is

² Ibid., pp. 28-29.

³ J. Rawls, *Political Liberalism*.

⁴ F. Tönnies, *Community and Civil Society*, J. Harris (ed.) (Cambridge: Cambridge University Press, 2001).

itself a Person, or in some important respects like a person.⁵ While initially based on legal fiction, later the *universitas* came to be understood as a unity derived from the substantive purpose that its members agreed to pursue jointly. As Oakeshott puts it:

It is not (or need we suppose it to be) many persons transformed in one Person. It is a many speaking with one voice, not because all the tongues have miraculously become one Tongue, or because all the intelligences concerned have become one Mind, or all the wills one Will, but because [...] they are agreed, not merely to speak the same language, but to say the same thing and are equipped with the means of committing themselves to or acknowledging such common utterances as their own.⁶

Something like this, the objection goes, is precisely what the fact of pluralism rules out. This kind of moral order no longer exists.⁷

What could be said in response to these three charges? I do not think it is fruitful to deny their empirical accuracy. The first and second claim are I think evidently true, and even though I am much less certain about how deep the fact of pluralism really goes, I shall here simply assume that it goes sufficiently deep to raise an important concern. Instead, what I want to do is to investigate whether the notion of shared intention can be adapted such that it can accommodate these concerns.

2. Collective Intention: An Interpretive Approach

Let us first investigate an argument that aims to render the corporate thesis more amenable to the complexity problem. Essentially, this argument is based on an idea advanced by Ronald Dworkin in his book *Law's Empire*.⁸ There Dworkin argues that we should think of intention in interpretive rather than psychological fashion. The

⁵ M. Oakeshott, *On Human Conduct* (Oxford: Oxford University Press, 1975) p. 203.

⁶ *Ibid.*, p. 205.

⁷ On a historical note, the fact and value of associative pluralism was the very premise of *Das Deutsche Genossenschaftsrecht* and the later work of British pluralists like Laski, Cole and Figgis. For the historical development of their ideas, see D. Runciman, *Pluralism and the Personality of the State*.

⁸ R. Dworkin, *Law's Empire* (Oxford: Hart Publishing, 1998), Chapter 2. His interpretive conception of intention goes back to the hermeneutic method pioneered by H. G. Gadamer, *Truth and Method*, second, revised edition (London: Sheed and Ward, 1989).

interpretive view departs from the commonplace assumption that an intention is a mental datum, a conscious mind-state or event that takes place before we act. It agrees that intention guides action, but it suggests, rather, that it is *implicit* in what we do. An intention merely manifests itself by the ‘constructive interpretation’ of practice or a sequence of actions.

Briefly, in the context of the (future-directed) intention of one individual, the gist of the interpretive idea is that we start from a temporal series of interconnected activities (or present aims) and then – in reverse, so to speak – try to construct the intention that ‘best fits’ this historical practice as a whole. So, the interpretive method purports to account for a sequence of actions $A_t, A_{t_1} \dots A_{t_n}$ such that the entire sequence could have been planned by one agent. Put differently, it seeks to capture what these actions added together aim to achieve, presumably at t_n . So the decisive issue here is not that this plan or aim figures as a conscious state in the mind of some person throughout the entire sequence. This may or may not be the case. Instead, the presence of a (future-directed) intention entirely depends on the question whether or not these actions (or present aims) do in fact cohere in this rational manner.

In the context of collective intention, the implications of the interpretive idea are similar. Again we start from a set of co-ordinated activities. But now these activities are not just connected in a temporal sequence; they are also connected in a bundle consisting out of the activities (or I-intentions) of a group of people $I_1, I_2 \dots I_n$. Here, the interpretive method aims to account for the entire practice as if it had been sanctioned by one rational agent. It tries to explicate how the activities of all group members cohere rationally, that is, how they collectively aim to achieve a certain end. Hence, a collective intention does not have to occur as conscious state in the minds of $I_1, I_2 \dots I_n$. This may or may not be the case. What is necessary, however, is that their activities (or I-intentions) are ‘meshing’ or ‘overlapping’.

In other words, I am invoking a notion of collective intention that does not necessarily correspond to what each participant in the practice actually meant by doing this or that. Some participants may just have meant to do as they were told, others may not have meant anything at all beyond what they were in fact doing. But constructing a collective intention is different from recovering actual mind-states. Above all, it aims to account for the rational purpose that is *immanent* in practice. In this way it gives a rational justification for what the participants do individually. While this justification does not necessarily sum up what they actually thought at the moment of acting – their subsidiary intentions – it does aim to capture how they, if pressed to account for their actions in terms of a reason, would ultimately express the purpose and meaning of their conduct.

On the interpretive view it is not necessary that all participants assume the executive perspective previous to taking part in collective action, as it is on Bratman's model. Such a demand would indeed cast a grave doubt over the corporate thesis in this chapter. As Dworkin candidly concedes in his interpretive theory of law, this holistic standpoint requires an intellectual effort of Herculean proportions. Within the resource constraints of their office usually even judges are incapable of making this effort. The same conclusion would appear to hold *a-fortiori* for most members of a political association. More drastically, it is conceivable that no human being could make that intellectual effort. Hayek seems to have thought it epistemically impossible for any individual or any central agency to assume the executive perspective even in the most ideal of circumstances.⁹ If true, the executive intention could be no more conscious a mind-state than Hegel's Cunning of Reason. On the interpretive view, however, it is not important that collective action is consciously designed. A configuration of second-tier intentions may well be the result of Spontaneous Order.¹⁰

⁹ See for example his *Law, Legislation, and Liberty*.

¹⁰ Note that Hayek also argued that 'the individual [...] must be ready and willing to adjust himself to changes and to submit to conventions which are not the result of intelligent design, whose justification in the particular instance may not be recognizable, and which to him will often appear unintelligible and

However, what *is* important is that the practice itself, that is, the subsidiary intentions that the participants do have, possess enough internal coherence to permit rational justification in terms of a collective intention. Be it designed or not, that practice must be ordered in a purposive or rule-based manner. In addition, what matters critically to the idea of corporate personality is that its collective intention is effective as a reason-giving *norm* or an *ideal* that constrains, judges and, if necessary, discards and modifies the associative practice itself. This effectiveness as a norm, in the end, makes the difference between a corporate person and a loose aggregate or arrangement of individuals.

Could an executive intention be truly action-guiding if it were merely immanent? Clearly, the standard presupposes the value of the historical practice or association *as a whole*, for it merely expresses ideas and principles that are already embodied and objectified in it. However, the interpretive idea does not presuppose that *each and every* act contained in the practice will also form the perfect embodiment of its idea or principle. Nor does it deny the possibility of critical reflection and change. We do reflect on what we do and we do so by appealing to abstract theory and principle. But on the interpretive view, theory and principle do not provide cosmic standards. They merely encapsulate the intention that holds the practice together intellectually and that gives it its integrity and internal justification. So, while we can appeal to theory, we cannot forget that theory is merely the intellectual residue of a way of life, a 'living morality'. And the upshot of this is that, intellectually, we can never advocate the *wholesale* rejection of that way of life. We can, however, find critical leverage by reflecting the practice as a whole into theory and then back into practice, claiming that some parts of the latter need adjustment to achieve a better fit with the former. In other words, we can say that certain actions or certain patterns of behaviour, belonging to a wider practice, ought to be judged by their own immanent or justificatory idea. And this

irrational.' See his 'Individualism: True and False', in: *Individualism and Economic Order* (London: Routledge and Kegan Paul, 1949), pp. 1-32, at p. 22.

critical attitude is precisely what corporate personality demands. For a group to qualify as a corporate person it must endeavour to make its decisions and actions cohere in a rational way, that is, their actions and decision must fit a collective intention. Of course, like individuals, groups may not always fully achieve this goal, but they must at least recognise that integrity is the standard to which they are held to account. They must acknowledge the normative pressure to adjust action to its justificatory idea. Perhaps this approach merely supports piecemeal reform, but caution and conservatism are not the enemy of effective criticism.

3. Two Unpromising Solutions

Thinking of intention in this interpretive fashion takes the edge off the complexity problem. So at least I argue. But does this interpretive turn also address the concerns about membership, boundaries and pluralism?

Taking the last concern first, note that it is still far from inconceivable that we end up with competing interpretations of a collective intention. As I understand the issue, the principal criterion for the success of an interpretation is that it should fit the practice in question. And an interpretation fits to the extent that it accounts for the goings-on in a practice as the collection of individual activities aiming to achieve one purpose or willed by one single agent.¹¹ By the lights of this criterion, the best conceivable interpretation will succeed in constructing an account showing how the practice in its entirety could be captured in the form of an executive intention. Evidently, this is an ambitious benchmark. It seems reasonable to settle for an interpretation that captures the greater part of goings on, giving special attention, moreover, to a range of paradigm cases. This is also how we could still gain a critical vantage point on a practice. The critic suggests that we reform the anomalous and tangential parts of the practice to its

¹¹ R. Dworkin, *Law's Empire*, pp. 65-68; 230-231.

underlying principle so as to attain an even greater fit. However, the trouble with this less ambitious criterion is that it leaves room for the possibility of irresolvable conflict. Interpretive justification comes to a halt, among others, when we disagree on which parts of the practice are central or paradigmatic and which peripheral.

Dworkin argues that to settle these conflicts we need to appeal to a second criterion that as far as I can see is not fully internal to practice. We may feel that a practice would altogether be of greater appeal morally or aesthetically if it were justified under one interpretation rather than another. So we have to prefer the intention that fits with the bulk of the data and also depicts the practice and its history in its 'best light'. But this is precisely where the interpretive approach may run up against the challenge of pluralism. After all, the pluralist points out that people disagree about such a criterion. And so we end up with different authoritative justifications and action-guiding principles.

In fact, pluralism seems to pose a dual problem for the idea of corporate personality so understood. If the state is to have integrity the following two conditions must *both* be met. In the first place, it is vital that the associative members arrive at a 'pre-interpretive' recognition of that what is to be interpreted: a practice composed of a certain set of activities.¹² Evidently, it only makes sense to construct a collective intention if it is clear what delimits one set of collective activities as one practice from another. Thus, as a minimum prerequisite, the citizens of a state need to have an *ex ante* understanding of what the state is or what it essentially does as opposed to other associations and collective practices. We have to demarcate a public realm, for example the activity that involves coercion backed up by law, before we can construct the purposes or principles that govern this activity by means of interpretation.

The second condition, as we already saw, points out that while a collective intention may not have to be a mental state or a thought in a mind, it requires a genuinely authoritative interpretation when brought to consciousness. We may allow some room

¹² Ibid., p. 66.

for interpretive differences, but it cannot suffice if people consistently differ about what they purport to achieve together. In other words, citizens should be able to recognise *ex post* that they share a particular intention, at least at the level where it is immanent in common practices or goings-on. We need to have access to a genuinely shared executive intention that can rule things in or out.

To sum up, the first condition demands that we draw the same boundaries around the historical practice that we think is constitutive of political association. The second condition requires a level of agreement on the executive principles that guide this particular practice. We must come to some post-interpretive agreement on the substance of a collective intention. Presumably, it is this last condition that causes the main problem. We need an executive intention that *all* consider to be authoritative within the public domain.

3.1 Liberal Nationalism

Here, then, the idea of corporate personality leaves no room for difference and pluralism. Many theorists who look favourably on community squirm when confronted with this upshot. Could we just piggy-ride on the back of the solutions they offer?

Consider David Miller, who in his book *On Nationality* tries to deal with this issue by making a distinction between conservative and good nationalism. He argues that the state should be implanted in national culture, custom and tradition because they provide the necessary settings for social justice, solidarity and democracy to flourish. But at the same time he baulks at 'conservative nationalism' for a number of reasons. Conservatism, he says, conceives of the nation 'not merely in terms of horizontal ties to fellow-members, past and present, who share whatever features are taken to constitute the common identity, but in terms of vertical ties to established institutions, which are regarded as authoritative.'¹³ The conservative view also involves, in Scruton's words, 'sanctity, intolerance, exclusion, and a sense that life's meaning depends upon

¹³ D. Miller, *On Nationality*, p. 124.

obedience, and also on vigilance against the enemy.’¹⁴ Hence, nationalism of this flavour promulgates hostility to those who do not share the national customs.

Miller makes clear that he believes these things to be wrong. However, his kind of nationalism ‘need not be (and as a matter of fact is not) authoritative in the way that the conservative supposes.’¹⁵ National tradition should not commend deference and it should not rely on the censure of criticism and exclusion. National values have no greater standing than the recipes in a cookery book, as he puts it at one point. As for non-nationals, ‘All it needs to ask [...] is a willingness to accept current political structures and to engage in a dialogue with the host community so that a new common identity can be forged.’¹⁶ Miller’s principal idea is here that the host’s identity must be ‘stripped’ of parts that are ‘repugnant’ to immigrants, and vice versa, with the purpose of blending these diverging identities into an all-inclusive national identity.¹⁷

Why not simply track the Miller solution? To begin with, I am sceptical that it offers a genuine way out of the problem. Miller’s language may sound vaguely benign, but if politics and the law are infused with and guided by standards of conduct that belong to the customs of the nation, what he writes here fudges the issue. If it is the nation-state that we need, then we cannot treat our national values as somehow on par with the ethos of others. Above all, we have to recognise that the national ethos is normative within the public domain. To be blunt, we have to maintain that our national identity determines what the state and its citizens should and should not be doing. And as the law demands compliance, conservatives are entirely right to frame the issue partly in terms of vertical ties, exclusion and obedience to authority. No honest nationalist could mean to say that the nation’s customs are like recipes in cookery books, which cooks may freely add to or simply ignore if they feel like it. Whatever else we want to say about these values

¹⁴ Cited in D. Miller, *On Nationality*, p. 125.

¹⁵ *Ibid.*, p. 129.

¹⁶ *Ibid.*, p. 130.

¹⁷ This, at least would appear to be his preferred view in *On Nationality*. For further discussion see also his *Market, State and Community* (Oxford: Clarendon Press, 1989), especially Chapter 9, and the essays in his *Citizenship and National Identity* (Cambridge: Polity Press, 2000).

and customs, at some level they must be authoritative. We may tolerate a plurality of practices in private life, but in public life there must be one overarching and one binding recipe or there cannot be a unified nation-state.

What is more, Miller does not seem to realise the full extent to which there are limits to how far identities can be stretched and moulded to assimilate difference into one all-inclusive identity.¹⁸ It is an undeniable fact, of course, that national identities are not cast in iron. What is deniable, however, is that change must consist in an indiscriminate process of scooping up all newly arrived identities and throwing them into a giant melting pot. The latter kind of change presumes that these identities are benign and compatible. But if pluralism is the problem, this may no longer be the case. Some substances do not mix. Moreover, stripping away the elements of the national identity that are repugnant to some groups may also seem a distinctly unappealing suggestion. Homosexuality, the consumption of alcohol and gender equality, to name a few possible examples, are practices that some cultures would probably reject as highly repugnant. British national values are based, among many other things, on tolerance, fundamental equality and perhaps no less on the consumption of copious amounts of alcohol, or so I imagine. Now, if Miller is right, not only do we need to welcome such perspectives to Britain, we should even be prepared to strip away these 'repugnant' elements and replace them by some half-baked compromise that would offend no-one. And to make matters still worse, this compromise would contain the standards that guide the state and thus claim final authority over its citizens in the public domain.

It is far from absurd to note that national traditions and identities change and evolve under whatever kinds of pressures are brought to bear on them, as conservatives are well aware.¹⁹ What is absurd is the suggestion that this fact somehow lends credibility

¹⁸ He acknowledges only two limitations in *On Nationality*, at pp. 128-129. The first limitation is that the influx of non-nationals cannot be allowed to take place too rapidly. The second that non-nationals may not continue to form a separate nation themselves.

¹⁹ In a footnote Miller acknowledges that conservatives know that traditions and identities are not frozen in time (pp. 126-128). Instead, he accuses conservatism of an inconsistency. Now, it may or may not be inconsistent to grant, on the one hand, that traditions do inevitably change over time and to reject any

to the view that we should never guard our national identity from these pressures. For that view presupposes that all change is to the good, or at least neutral, and all resistance to change bigoted prejudice. What is perhaps even more absurd is the idea that it is in fact possible to amalgamate all different viewpoints into one national identity that would still make available to us the kind of spiritual home that nations are thought to offer. I have a sneaking suspicion that there are limits as to how much the melting pot can process even in theory, let alone in practice. But putting this issue aside, if a national identity were really as absorbent or spongy as Miller here claims, what good would it still do? Presumably, any collective identity must at least give us a strong sense of who we are and what we stand for. In Miller's case it must also generate a social context for trans-generational solidarity, social justice, democracy and a host of other splendid things he and other egalitarians are partial to. For example, he claims: 'Because our forebears have toiled and spilt their blood to build and defend the nation, we who are born into it inherit an obligation to continue their work, which we discharge partly towards our contemporaries and partly towards our descendents.'²⁰ Elsewhere he further points to the fact that 'we hold nations with their present memberships morally responsible for the deeds of their forebears; for instance we have recently witnessed the Queen apologising on behalf of Britain to the Maori's of New Zealand for injustices perpetrated in the nineteenth century'.²¹ However, if this is the mushy stuff that national identity is made of, then I fail to see why Miller expects it to contribute decisively to the realisation of these ends. Miller is I think unduly prepared to cave in under the pressure of multiculturalist criticism. Much as he would like to escape the whiff of conservatism, that view comes much closer to what a viable defence of the nation would actually look like than the attempt to steer a middle course between nationalism and multiculturalism.

form of change on the other. But surely it is not inconsistent to combine that fact with the claim that some changes are more welcome than others or the view that we ought to salvage what is good in the past and reject what is bad in the future.

²⁰ D. Miller, *On Nationality*, p. 23.

²¹ D. Miller, 'Nationality: Some Replies', *The Journal of Applied Philosophy*, 14 (1997), pp. 69-82 at p. 78.

3.2 Liberal Impartiality

We have to look for an alternative to liberal nationalism. Could we perhaps find it in liberal impartiality? This view argues that states should be guided by justificatory principles that no reasonable and suitably motivated person could possibly reject.²² When based on such principles, states would be neutral to the plurality of reasonable, local identities that we may come across in contemporary societies. Instead, they would invoke the principles of ‘the party of human kind’, as Korsgaard puts it. This works out in such a way that states could only aim to facilitate or redistribute certain indispensable resources or goods that enable its citizens to pursue the purposes and life-plans of their own choosing. Here, I do not want to discuss a particular conception of what these indispensable goods are. And I certainly do not want to commit myself to one distributive rule or another. The bottom line is that liberal impartiality claims to rise above the fray of diverging ‘comprehensive doctrines’ by appealing to what it thinks are universal rather than parochial values. To put it in terms of Rawlsian political liberalism, the guiding and justificatory principles in the public domain capture the spirit of an ‘overlapping consensus’.²³

This is not the place to discuss the prospects of liberal impartiality at length. However, there are at least two grounds for general caution with this solution. First, while impartial principles of justice may plausibly specify the content of the collective intention of the state, it is still imperative that it be shared at the societal level, that is to say, these principles must be actually recognisable as normative standards to all citizens. Now, as I understand it, the liberal impartiality argument claims that all reasonable and suitably motivated persons *would* agree to its principles. However, this does not mean that people *will* in fact so agree. The contractualist test is of course entirely hypothetical and relies on a value-laden conception of ‘reasonableness’, which

²² For example, see B. Barry, *Justice and Impartiality* (Oxford: Clarendon Press, 1995).

²³ J. Rawls, *Political Liberalism*.

seems to fix certain outcomes in advance. Thus, the strategy's success depends in part on how deep the fact of pluralism actually runs. If the fact of pluralism extends to our understanding of reasonableness itself, then little would be gained.

Second, even if pluralism runs less deep, the idea that we will *ipso facto* finish up with polities whose boundaries are congruent with those of present states is far too neat to be plausible. This leads to the problem of membership. If integrity could only be defined in terms of increasingly universal and impartial principles such as liberty, equality or fairness, then we cannot credibly expect to end up with the concrete identity and membership of a territorially bounded state. Certainly, it may be true that the collective intention of the British State is in part shaped by these values and ideals, but it is surely not these notions that set it apart from the French State or the United States. The problem is that such abstract ideals are likely to be shared by a public much wider than British citizens alone. If there is some shared identity of this sort, it will not correspond to a national or even political identity but to a moral community perhaps impossible to delimit geographically. The great difficulty with the personality of the state is not just that it needs to accommodate a degree of pluralism; it also needs to do justice to the particular and physical manifestation that is the state. States are formed by precisely delineated groups. Moreover, states have jurisdictions that are defined in equally precise geographical terms. So even if we stumble upon some shared identity at this level of abstraction, it still seems highly improbable that this could be cashed out in ways that correspond to the actual boundaries and the actual membership of states.

4. Institutional Persons

If pluralism is indeed a fact, the interpretive account draws its boundaries either too narrowly or too broadly. In order to address this problem, we therefore need to add a further ingredient to our account of state personality. The remainder of this chapter

argues that we need to extend this notion of state personality with an institutional dimension. Thus far, I have assumed that integrity is a purely deliberative and egalitarian notion. On this account, the notion of state personality resembles the Rousseauian ideal of the psychological conversion or the melting of individual wills into the General Will of the state. This conversion takes place on the two-tiered basis of a collective intention, a plan or principle that is action-guiding for the group *as a whole*, and subsidiary intentions and objectives that guide the activity of the participants *as individual*. So the person of the state emerges if this collective intention is effective as an action-guiding principle, that is, if its citizens *directly* come to accept the collective intention as the standard by which their (relevant) individual desires and actions are rationally constrained.

However, I want to distinguish this *deliberative* account from an *institutional* approach. The latter approach suggests that the effectiveness of a collective intention could also depend on the acceptance of certain intermediary mechanisms. Essentially, these mechanisms coin procedures for the identification of particular expressions of intent and particular decisions as the expressions and decisions of the state. So they aim to establish certain criteria that amidst the cacophony of voices recognise which voice counts as the voice representative of the association. Hart, in his book *The Concept of Law*, famously refers to such intermediary mechanisms as ‘rules of recognition’.²⁴ To name a few examples: rules become binding law only when ratified by ‘Queen in Parliament’; individual transgressions of the law are adjudicated decisively only in courts of law, before the appropriate magistrates; the ‘Will of the People’ is only identified in a particular ballot governed by certain counting methods and other procedures. To put it another way, Hart’s rules of recognition are formulae that transform certain individual acts and decisions into acts and decisions of the association. They give the state a human face, for example, that of a King, President or Prime

²⁴ H. L. A. Hart, *The Concept of Law* (Oxford: Clarendon Press, 1997), second edition, p. 94 ff.

Minister. On a purely institutional account, the rules of recognition are constitutive of the state and its personality. Without them there could be no Law, no Adjudication, no People, no unified action, but merely the chaos of conflicting intentions each claiming to be authoritative for the whole.

To clarify the general nature of these rules somewhat further, it may be useful to briefly consider Searle's theory of 'constitutive rules'.²⁵ In the spirit of Hart, he defines constitutive rules as rules that, rather than regulate a pre-existing practice, create an entirely new kind of fact or activity. One of his examples is the game of chess. Searle points out that the laws of chess could not have been invented subsequent to the activity of chess-playing since they delineate what the game is about in the first place. They specify that certain bits of wood count as 'King', others as 'Rook', again others as 'Bishop'. And so they turn dumb bits of wood into chess pieces with a particular function and significance. They further declare that you win the game by checkmating your opponent's King, that Bishops may move in diagonal direction only, and that 'Pg7-g8(Q)' means the promotion of Pawn to Queen. Of course, it is possible to fiddle with dumb bits of wood on a chequered board, but a purely physical activity like this could hardly be described as chess. It becomes chess-playing only after the rules have been put in place.

The game of chess is merely one example of what Searle describes as the broader category of 'institutional facts'. In general terms, all institutional facts are 'two-dimensional'.²⁶ On the one dimension there is a 'brute' event X, some purely physical state of affairs or goings-on. (The fiddling around with dumb bits of wood.) On the other dimension there is an institutional matrix that assigns a particular meaning or status or function Y to this brute event in a particular context C. (The laws of chess.) Formally, as Searle points out, an institutional rule states that 'X counts as Y in C'. On

²⁵ J. R. Searle, *The Construction of Social Reality*, especially pp. 27-29.

²⁶ As Taylor puts it in his 'Irreducibly Social Goods', in: *Philosophical Arguments* (Cambridge MA: Harvard University Press, 1995), pp. 127-145, at p. 134.

the view that I am entertaining here, states too are institutional facts. That is, they are based on rules that stipulate that X counts as Y in C. For example, the fact that Britain has a Queen has a brute, physical aspect. The Queen, presumably, is a human being like everyone else. But this fact also depends on the rule that identifies her as 'Elizabeth the Second, by the Grace of God, of the United Kingdom of Great Britain', and so forth. The fact that Britain has a People has a similarly brute aspect. But it also relies on rules that identify the Will of the British People through elections. Strip away the principle of majoritarianism and what do you have?

With this notion of rules in hand, let us now return to the main problem. The institutional model claims that the personality of the state is mediated by rules of recognition or constitutive rules. What does this entail? Essentially, I claim that it shifts the *locus of integrity* to a different level. To recapitulate, the general idea of corporate personality is premised on the notion of a collective intention. I already argued (in Section 2) that this intention is internal to a practice, that is to say, immanent in certain overlapping intentions and activities rather than a conscious mind-state itself. But even so, it is still effective as a normative constraint on that practice and the subsidiary intentions that it involves. Now, the idea of a locus introduces another element. It specifies *where* that collective intention should be effective as a normative constraint. And since this intention is supposed to be internal to practice, it also specifies where that practice is situated. I argue that the locus is either *deliberative* or *institutional*. I now proceed by unpacking this distinction.

The Deliberative Model essentially demands that the population at large accepts one collective intention as authoritative. Since intention is read here in interpretive rather than psychological fashion, this implies in effect that we must all be engaged in the same social practice so that the collective intention internal to that practice is normative for us. The model also supposes the absence of interpretive conflict when this intention is brought to consciousness. But the main requirement is that there is a coherent practice

that we all participate in. It is precisely there that the collective intention should be effective as a normative constraint. For example, what we would need is that the members of the state affirm, merely by doing what they normally do, the normativity of a common good. Or they must in the same implicit way affirm the normativity of a collective intention working towards the achievement of a conception of justice.

Now consider the Institutional Model. This model suggests that once certain procedures or official bodies are in place to legislate, adjudicate and govern on behalf of the community, the locus of integrity shifts from the population at large to the state's institutions. While its citizens no longer have to acknowledge the normativity of an executive intention – they just need to obey – the virtue of integrity applies in full force to the legislative, policy-making and judicial institutions of the state. These institutional responsibilities must be guided and constrained by an executive intention, that is, by the justificatory principles that account for the state's decisions and actions as those of one single and continuous agent.

Essentially, the model submits two closely related claims that need to be addressed separately. As I just mentioned, it first suggests that the collective intention may be effective at the institutional rather than deliberative level. So the norm of integrity applies to *institutions* rather than *individuals*. I shall come back to this argument a bit later because it already presupposes a second claim. While it is not necessary that we all recognise the direct authority of an executive intention, this claim argues that the rule of recognition, that is, the rule that allows certain officials to speak on behalf of the association, needs to be widely accepted. I look at this last claim first.

4.1 The Social Acceptance of Institutions

Hart makes an important distinction between 'primary rules' and 'secondary rules'.²⁷ The former focus on 'the actions that individuals must and must not do'; the latter

²⁷ H. L. A. Hart, *The Concept of Law*, pp. 91-99.

concern the primary rules themselves: 'They specify the ways in which the primary rules may be conclusively ascertained, introduced, eliminated, varied, and the fact of their violation conclusively determined.'²⁸ While primary rules do not have to be shared on the Institutional Model, it is absolutely critical that the secondary rules are widely accepted. According to Hart, acceptance involves two things.

First, it means that these rules are firmly entrenched in the patterns of conduct of the members of the association. It is possible to codify rules of recognition, as often happens, but formal codification is never quite sufficient. To put it differently, the constitutive rules of an association cannot be fixed or put in place by some external source. This would merely beg the question of the authority of this source. Instead, they depend entirely on the effectiveness and the public knowledge of the rule. The idea is that the rules in question are not imposed on a practice, but merely describe what is the normal conduct of its practitioners. They are customary rules or rules in practice. According to Hart, if judges, lawyers and civil servants generally treat decisions from Parliament as legislative, then this is the accepted rule of recognition for law.

Searle's example of money could be used to emphasise that constitutive rules are rules in practice.²⁹ In our current monetary systems, coins, banknotes and the like acquire their function as tender not because of their intrinsic value, which is of course slight, but only on the basis of the fact that enough people recognise and accept these objects as a form of tender. Usually all this happens on the basis of tacit assumption, but when I exchange a couple of coins for a cup of coffee, I do so on the basis of a collective acceptance of the rule according to which these metal things count as tender. Should sufficient people revoke this acceptance, which is what happens when confidence in a currency is lost, only the brute fact of a heap of metal and paper remains. Governments may attempt to shore up what they claim is legal tender, but the survival of a currency ultimately depend on our own perceptions.

²⁸ Ibid., p. 94.

²⁹ See the discussion of this example in *The Construction of Social Reality*, pp. 37-51.

Note that institutional facts still rely on collective intention. The difference with the Deliberative Model lies in the content of the intention. To use Searle's taxonomy, collective intentions divide into 'social facts' and 'institutional facts'. He speaks of social facts when the content of collective intention refers to some activity we undertake together, for example, going for a walk. On the institutional notion, however, the content of a collective intention merely consists in assigning a functional status to brute or social facts.

It is also important to point out that constitutive rules tend to be second nature. In everyday life many institutional facts are treated precisely like brute facts. We do not normally have to tell long stories explaining, for example, why the Queen is Queen or why the Prime Minister is Prime Minister. To most of us it is perfectly plain that I would be both an impostor and a fool if I challenged these facts by claiming to be the only true King myself. These are facts that we 'just know' are true, like chess players 'just know' that Pawns are Pawns. But we only know them to be true because they depend on an institutional background so firmly implanted in practice that they do not even have to be made explicit. Like chess pieces on a board, Kings and Queens and Prime Ministers exist by the mercy of a collectively accepted set of secondary rules.

Second, like primary rules in the Deliberative Model, the secondary rules must be endorsed from what Hart calls the 'internal point of view'. From this point of view, rules are not just descriptive or predictive of what people are up to, but also normative. That is to say, they are taken to provide standards for our own present and future conduct. It is important to stress that the internal point of view does not imply that rules in practice have to be just, moral or democratic. Nor does it imply that they have to be based on an express consent. What it means is that they have to be affirmed as reason-giving. Secondary rules do not only illustrate how primary rules come about in a certain state, but also how they *ought* to come about.³⁰ As Hart points out, the normative basis

³⁰ *The Concept of Law*, pp. 88-91. Here I should stop to point out that I depart from Hart's view in the following respect. Hart suggests that a legal system exists once the courts or the officials who in one way

of the rules of recognition might be one of moral conviction. But it might also just be a matter of convention or respecting tradition. In the latter case, the rules are reason-giving or regulative of future conduct simply because they are and have been widely held and respected.³¹

4.2 *Why Integrity Still Matters at the Institutional Level*

This does not imply, however, that the norm of intention has no impact at the institutional level. As I understand the issue, the emphasis on secondary rules is not incompatible with the interpretive emphasis on collective intention. The Institutional Model just shifts the normative jurisdiction of intention to another locus. Put differently, while the general acceptance of a rule of recognition is a *necessary* condition of institutional personality, it is not in itself a *sufficient* condition. In addition, these institutions must be bound by the normativity of an executive intention.

There is, of course, a brand of legal positivism, which is commonly associated with Hart, that claims that we do not need collective intentions at the institutional level. Briefly, this view holds that the rules of recognition form the only source of the law. Once these rules are clearly specified and widely accepted, the courts know what the primary rules are and merely have to apply them. We have to turn to Dworkin's conception of 'law as integrity' for a view on which the notion of an executive intention continues to matter.

or another practice law accept and assume the internal perspective of secondary rules. See *The Concept of Law*, pp. 113-117. It is not necessary that ordinary citizens take the same perspective. Citizens 'only' need to obey, out of habit, fear or for whatever reason. They may be 'deplorably sheeplike', as he puts it (p. 117). However, I fail to see the rationale for this particular division of labour. It is true that citizens could be 'sheeplike'. One of the features of the institutional model is that we only have to obey *primary* rules and not also regard them as reasons. But in my opinion 'sheep-likeness' is not enough in the case of *secondary* rules. It is evident that the practitioners of law have to accept the rules of recognition, but why should the law have any meaning for us, the subjects of law, if it is self-proclaimed? The idea that citizens grasp the full complexity of secondary rules is perhaps a fiction, but surely it is not fictitious to suggest that citizens accept their most basic institutions such as Parliament and the courts.

³¹ For this claim, see Hart's 'Postscript' in *The Concept of Law*, P. A. Bullock and J. Raz (eds.), pp. 254-259.

According to Dworkin, jurisprudence is at least partially interpretive. Acceptance of secondary rules does not suffice on his view. In order to reach the right decision, particularly in so-called 'hard cases', the courts have to construct and appeal to justificatory principles that underpin the practice of law as a whole. And these principles are obtained by means of the interpretive method, that is, by accounting for the historical practice of law as the legal intent of one single agent. Not until these principles are taken as a source of law does it become fully clear what the law really demands, or so his theory goes. However, while Dworkin argues that we need interpretation and justificatory principles to identify the law, he has to admit that we need to know what the practice of law is *before* interpretation can lead us towards these principles.³² So he cannot dismiss the importance of institutional rules either. 'Law as integrity' needs to have a social source that is independent of principle, that is, it needs socially accepted rules of recognition.³³ In effect, this would suggest that Hart's rules of recognition and Dworkin's emphasis on intention are *jointly necessary*. While the general acceptance of a secondary rule is indispensable, the law hinges as much on its integrity.

My point is here not that this or that theory of law is the most persuasive. It is rather that Dworkin's account of institutional integrity describes what we mean when we apply the notion of collective intention or personality at the level of the state. What does collective intention or integrity stand for? Essentially, it demands of institutions precisely what it demands of individuals or deliberative groups.

We want our neighbours to behave, in their day-to-day dealings with us, in the way we think right. But we know that people disagree to some extent about the right principles of behavior, so we distinguish that requirement from the

³² Dworkin writes that for someone to interpret something, 'He needs assumptions or convictions about what counts as part of the practice in order to define the raw data of his interpretation at the preinterpretive stage; the interpretive attitude cannot survive unless members of the same interpretive community share at least roughly the same assumptions about this.' In: *Law's Empire*, p. 66.

³³ This is part of Hart's critique of Dworkin in his 'Postscript'. It may or may not be necessary to construct Dworkinian principles to fully identify the law, but the objection is that such principles are themselves not sufficient either.

different (and weaker) requirement that they act in important matters with integrity, that is, according to convictions that inform and shape their lives as a whole, rather than capriciously or whimsically.³⁴

In other words, their aims and actions have to be authoritatively governed by an intention that justifies all their actions as one cohering, rational whole. Institutional decision-making should be subject to the same standard. As Dworkin puts it, the state should be 'faithful to its own principles'.

Dworkin is mainly interested in jurisprudence. Law as integrity rules out legal decisions that rest on a compromise or inconsistency at the level of fundamental principle. To borrow one of his examples, consider a law that allows abortion in randomised fashion. Say that it sanctions abortion for women born in even years and rejects abortion for women born in odd ones. Assume that this happens to provide a *modus vivendi* in a society deeply divided on the issue. Assume further that this 'checkerboard law' has travelled through all the appropriate legislative institutions. Thus, it is based on past decisions of the right kind. While for the legal positivist this fact must be decisive, integrity in jurisprudence still blocks this law. For it demands that we are capable of explaining all decisions of the courts as the decisions of one unified and rational agent.

Notice two further things. First, the virtue of institutional *integrity* remains distinct from institutional *morality*. Kantians like Korsgaard put it to us that intention requires a person to act and decide in accordance with higher order principles that he could reasonably want to guide, rank and constrain the aims of all persons.³⁵ But I do not think we are logically bound to adhere to the strictures of impartiality to qualify as intentional agents, and I maintain that the same applies to states. Integrity is a non-moral virtue distinct from justice. It merely claims that individuals and states need to act in accordance with executive principles that guide, rank and constrain their subsidiary

³⁴ R. Dworkin, *Law's Empire*, p. 166.

³⁵ See my discussion in Chapter 2, Section 5, above.

ends. The claim, then, is that institutional decision-making is normatively bound by the collective intention that accounts for the historical practice of law *as a whole*.

Second, the normativity of collective intention does not just apply to judicial decisions. It directs the much broader process of legislation and policy-making. In this wider institutional context, one view that parallels legal positivism – call it democratic positivism – argues that once a certain institutional rule is accepted, legislation or policy will simply follow suit, without being subjected to further normative tests. To take the bull by the horns, assume that the relevant rule of recognition is formed by a particular majoritarian voting procedure. According to democratic positivism, we have no choice but to embrace whatever legislation or policy that procedure generates. However, if we want to advance a view of the state as a corporate person, we also need to reject this view.

Phillip Pettit has recently clearly showed why.³⁶ He argues that purposive associations in general face what he describes as ‘the discursive dilemma’:

Any groups that seek to make deliberative, reasoned judgments [...] face a dilemma. They may maximise responsiveness to individual views, running the risk of collectively endorsing inconsistent sets of propositions. Or they may impose the discipline of reason at the collective level, running the risk of collectively endorsing a conclusion that a majority of them – perhaps even all of them – individually reject.³⁷

To clarify the argument, consider one of his examples.³⁸ Assume that a group needs to decide on a syllogistic argument, which has as its first premise ‘*P*’, its second premise ‘*if P, then Q*’ and as its conclusion ‘*Q*’. Further assume that this group conventionally arrives at a collective standpoint by counting the votes of individuals A, B and C, and

³⁶ Pettit sets out his argument in a series of recent papers: ‘Deliberative Democracy and the Discursive Dilemma’ (2001); ‘Groups with Minds of their Own’ (2001) and ‘Akrasia, Collective and Individual’ (2001), *ANU Working Papers*, all available at <http://socpol.anu.edu.au>. See also his *A Theory of Freedom*, Chapter 5; C. List and P. Pettit, ‘Aggregating Sets of Judgments: An Impossibility Result’, *Economics and Philosophy*, 18 (2002), pp. 89-110.

³⁷ P. Pettit, ‘Groups with Minds of their Own’, p. 10.

³⁸ P. Pettit, ‘Akrasia, Collective and Individual’, p. 7.

let the voting procedure be majoritarian. Finally, suppose that A, B and C hold the following opinions on this matter:

	<i>P</i>	<i>If P, then Q</i>	<i>Q</i>
A	Yes	No	No
B	No	Yes	No
C	Yes	Yes	Yes

Table 1. *The Discursive Dilemma I*. Source: P. Pettit, 'Akrasia, Collective and Individual', p. 7.

The group could make its decision on the basis of its widely accepted rule of recognition. However, what happens if the group is 'maximally responsive' to the views of A, B and C? Pettit points out that if we simply count their votes on each of these propositions, the resulting collective standpoint will be incoherent. The group believes '*P*', it also believes '*if P, then Q*', yet it rejects the conclusion that '*Q*'. Note that individually the views of A, B and C are logically perfectly valid or coherent. Thus, the dilemma is not caused by a defect attributable to the reasoning of the individual voters themselves. The problem is purely due to its decision rule.

Alternatively, we could, as Pettit puts it, 'impose the discipline of reason at the collective level'. This decision method prioritises the internal consistency of the collective standpoint. But it inevitably entails that the majority position on at least one of these propositions is sacrificed. Pettit further claims that the most obvious (but not only) way of doing this is to allow a vote to first decide on both premises and then to *deduce* the logically inevitable conclusion without further regard. On this 'premise-driven' approach, the group endorses *Q*, even though there is no majority for this conclusion.

Consider a second of Pettit's examples.³⁹ A political party has to decide on what platform to run. It does so in three temporal stages. It first needs to reject or embrace an

³⁹ I have added a minor adaptation.

increase in tax in March, knowing that a negative decision will allow for an increase of spending on either health or defence, but not on both. The party then needs to decide whether or not to advocate an increase in government health spending in June. Lastly, it has to decide whether or not to advocate an increase in defence spending in September. Assume further that it is a widely accepted rule in the party that its decisions are based on the votes of individuals A, B and C, and suppose that they have the following preferences.

	<i>March</i>	<i>June</i>	<i>September</i>
	Increase Taxes	Increase Health Spending	Increase Defence Spending
A	No	Yes	No
B	No	No	Yes
C	Yes	Yes	Yes

Table 2. The Discursive Dilemma II. Source: P. Pettit, 'Groups with Minds of their Own', p. 13.

Again, individually, the preferences of A, B and C are coherent. But if the party wants to decide all three issues by majority vote, it will clearly struggle to explain the rationality of its political programme. Pettit thinks it is palpably evident that no serious organisation could tolerate this kind of irrationality, and he is probably right to remark that this party would quickly become a 'laughing-stock'. Again, the alternative is to maintain the rationality of its proposals at the collective level. But this means that the party will either have to reject the vote in September or renege on earlier commitments. What is more, it is highly unlikely that the party could consistently renege on one of its earlier commitments. It could do so occasionally, perhaps, but it would clearly lose all its credibility if it reconsidered too often.

The idea of personal integrity implies that these kinds of considerations hold similarly for legislatures and governments. That is to say, there is a normative pressure 'to impose the discipline of reason at a collective level'. Pettit believes that this pressure

is traditionally captured by republicanism and the idea of deliberative democracy.⁴⁰ I take it to indicate that states could and should be organised as corporate persons. To say that a state is a corporate person, on this view, is to say that it does not rely fully on its majoritarian institutions. It is to point out, also, that it is not even sufficient to rely on the individual rationality of its voters. To ascribe personality to states, or any other association for that matter, their legislative and governing bodies need to be placed under the authority of a collective intention. So not only do we need the social acceptance of secondary rules, these governing institutions also need to generate legislation and policies that are justifiable in terms of the rational coherence of one unified and continuous intention.

4.3 Pluralism and Boundaries

Why is the Institutional Model better placed to get a grip on the issue of pluralism? The nub of the matter is this: If the secondary rules are relatively uncontested, it is strictly inconsequential that there is widespread disagreement about the content of primary rules. Ultimately, the executive intention of the state emerges not by reference to the claim that primary rules are reason-giving for all, as they would on the Deliberative Model, but by the general acknowledgement of certain institutional procedures that prescribe how the primary rules are created, changed, executed, and so forth. Integrity is only applied as a norm at the institutional level of the courts and legislative bodies of the state. Thus, citizens do not have to mesh their intentions as Bratman's model suggests and they do not even have to agree; they only have to obey.⁴¹

We could expound this claim with help of the interpretive conditions that were introduced above in Section 2. There I suggested that *both* conditions had to be met. The first condition demands that a practice is widely recognised *ex ante* as one type of

⁴⁰ P. Pettit, 'Deliberative Democracy and the Discursive Dilemma'.

⁴¹ To avoid any misunderstanding, the idea of institutional integrity only *allows* pluralism to exist. It does not, of course, offer any *assurances* that the state will in fact respect pluralism, reasonable or any other kind. And it certainly does not claim that the state *should* respect pluralism.

activity. It demands that we share certain pre-interpretive assumptions about what the state is or does. The second condition requires widespread, *ex post* agreement on the content of the executive intention within this realm. But while this is true on the Deliberative Model, the Institutional Model only needs the first condition to be met. If we take as a working assumption Dworkin's view that the relevant practice is delineated by the practice of law, and if law is 'the scheme of rights and responsibilities [...] that license coercion because they flow from past decisions of the right sort', then in order to meet this condition, we would merely have to agree on what decisions are 'of the right sort'.⁴² To share such an understanding is to accept an institutional matrix or a secondary rule that transforms certain 'brute' decisions into decisions that 'count as' law in a particular context.⁴³ There is, as we have just seen, a third condition, namely that institutional decision-making needs to acknowledge the normativity of a collective intention. But this condition is not affected by the fact of pluralism. Rather than requiring a societal consensus at the level of principle, it leaves this dispute for these institutions to settle.

Some additional points need to be made to get the Institutional Model up and running. To begin with, it could be objected that the fact of pluralism may not be restricted to conceptions of the good and the right. The rules of recognition could be hotly disputed themselves. If such disputes are rife it is no longer obviously true or false for someone to claim to speak on behalf of the association. Furthermore, like conceptions of the good and the right, institutions are subject to change. Now, I readily concede that there are limits to how much pluralism and change the idea of corporate personality could bear. In fact, these limits are a direct consequence of having defined corporate personality in terms of shared intention and integrity rather than fiction or spiritual substance. While fiction or Selves could not be touched by empirical fact, the

⁴² R. Dworkin, *Law's Empire*, p. 93.

⁴³ Dworkin uses the distinction between 'concept' and 'conception' to make a similar point. See *Law's Empire*, pp. 68-76.

idea of shared intention will always call for some rule or intention to be shared at the societal level. However, to me this seems to be much less of a problem for two reasons.

First, on a more general note pertaining to institutional change, the idea of the personality of the associations rests clearly on a Humean conception of personal identity over time that admits to degrees and change.⁴⁴ As we have seen, it is no anomaly on this conception that persons change if this happens gradually. On the Deliberative Model, the personal identity of a state is consistent with the fact that its political values and principles gradually change. Similarly, the institutional conception of corporate personality advanced here is not at odds with institutional change *per se*. There may be institutional change that implies that we can no longer sensibly talk of the continued personality of a state, for example, in the case of revolution, secession, civil war, annexation, and so on. But insofar as it tends to be gradual and sluggish, institutional change posits no major obstacle to the thesis advanced here.

Second, the claim that certain basic rules of recognition are widely shared seems empirically less vulnerable to pluralism than the claim that we share an entire political morality. It may or may not be true that monolithic, societal cultures no longer exist in the West, but it is evident that states still do. Most pluralists, presumably, would not contend that the fact of pluralism has advanced to the stage where states themselves are the chimera of a lost age of law and order. It is one thing to argue that Western societies like Britain, France or the United States harbour a great many conceptions of the good and even of justice, but it is quite something else to claim that they do no longer have an institutional framework that enjoys a basic degree of legitimacy and stability. Whatever kind of moral or cultural pluralism they may have to face up to, these and many other states do not teeter on the brink of revolution, civil war or annihilation.

Without denying its limitations, I simply press this model because its institutional fabric provides an added stability and continuity that is perhaps less likely to be derived

⁴⁴ See Chapter I above.

from the deliberative account. Consider again what David Miller is driving at in his book *On Nationality*. Although I am sympathetic to his view, pluralist critics may – perhaps not without reason – express doubts as to whether the British nation, as Miller sees it, has enough integrity and continuity to merit such conclusions. However, such doubts are perhaps less obviously warranted in the case of the institution of the state. Few people I think would question the more concrete continuity of Parliament.

In addition, the institutional model seems much better suited to do justice to the issue of territorial boundaries and membership. The Deliberative Model, even if it advanced principles of impartial justice, would still have to show that its ethos is somehow confined to a particular population, a specific set of practices, within a precisely defined geographical context. While I find such an argument difficult to conceive, geographical contexts do not look out of place within the institutional approach. It is far from inconceivable that the secondary rules assign authority to certain people or bodies of people within a physical jurisdiction or domain. Incorporating such a territorial domain requires a fairly straightforward extension of the model. For example, the rules of recognition could specify that decisions of ‘Queen in Parliament’ count as law, but only in the specific geographical context defined as the territorial boundaries of Britain. On reflection, it seems fairly obvious that the territorially boundaries of states are such institutional facts. Certainly, frontiers exist in a brute way, and physically their location may be more or less natural, but their institutional significance as the border of a state derives solely from a widely shared rule according to which certain geographical lines count as the boundaries of the jurisdiction of a state.

What is more, a similar extension would seem to take care of the membership issue. Bratman’s notion of shared intention suggests a membership criterion that is exclusively attitudinal. Whether I am an insider or outsider to a particular group depends solely on whether or not its members and I have the right subsidiary intentions and acknowledge the normativity of an executive intention. Clearly, this is not how states operate. I

cannot become a citizen of the United States just by subscribing to the principles underpinning its constitution. There is a myriad of additional conditions that need to be met. One of these conditions stipulates that I have to underwrite the constitutional principles of the United States, but this does not even begin to sum up what needs to be done. The Institutional Model has no difficulty in dealing with such conditions. It suggests that membership rests on institutional arrangements. The constitutive rules of a state do not only determine what counts as law, they also decide who counts as citizen.

Finally, I am not suggesting that the Institutional Model exists in a pure form. Theoretically, the corporate personality of the state is possible even when *no one* agrees to its primary rules. But in practice it is not very likely that a state could persevere – let alone thrive – without such support. That said, there is no reason to suggest that institutional integrity cannot be combined with deliberative elements. Only to the extent that a society falls short of the deliberative vision, does integrity depend on representative institutions and the attentiveness of these institutions to integrity. The orthodoxy is that scale is the decisive factor here. As the scale of an association diminishes, deliberative integrity gains in credibility. However, as this scale increases, corporate personality tends to become more reliant on intermediary institutions. The individual is perhaps the paradigmatic example of an association with deliberative integrity. The corporate personality of the state, it has been my claim, depends more heavily on an institutional chassis.

Conclusion

This chapter aims to extend the concept of corporate personality from two and three person examples to associations of a much larger scale including the state. Three problems appear to block such an extension. First, as political associations are often extremely complex, we cannot realistically expect its members to consciously assume the executive perspective. Second, while the notion of shared intention suggests a

conception of group membership that is relatively 'open', state membership is more narrowly defined. States, moreover, are confined to a certain geographical space. Third, the corporate thesis seems to rest on a deeply communitarian view of society. It presupposes moral consensus at the societal level that cannot be reconciled with the fact of pluralism.

I set out to meet these challenges with two different arguments. First, I offered an account of intention that is interpretive rather than psychological. This account departs from the view that collective intentions are mental data or conscious states in the minds of the group's members. While it agrees that collective intention must be action-guiding, it also suggests that it tends to be implicit in a practice. So we no longer have to suppose, somewhat naïvely at this scale of complexity, that associative members assume the executive perspective. It is adequate if their subsidiary intentions are sufficiently meshing in order to be able to construct the collective intention, the guiding plan or principle that fits the practice as a whole. Corporate personality, moreover, depends on the effectiveness of this executive intention as a normative constraint.

Second, I argued for a shift in the locus of integrity from the individual to the institutional level. Such a shift is possible on the basis of the general acceptance of a set of rules of recognition. These rules further limit the need for meshing subsidiary intentions to the institutional practices of legislation, policy-making and adjudication. So it is only the decisions and actions of institutions that need to be meshing in a rational way. It is only there that integrity applies as a normative standard. Hence, within limitations the institutional approach allows for greater social pluralism. The notion of institutional rules also allows us to come to grips with the problems of boundaries and membership.

I return now to the overall thesis. Together, Chapter III and IV develop a concept of corporate personality that I argue meets three objectives. First, by framing corporate personality in terms of shared and collective intentions we remain faithful to

reductionism about persons and personal identity. Second, by pressing an interpretation on which of shared intention remains independent from ethical individualism, we have opened up a conceptual space for corporate personality to do normative work. Third, by injecting interpretive and institutional elements into the basic model, we can now see how corporate personality could be applied to larger and more complex associations including the state.

This leaves a fourth objective. Chapter III essentially argues that integrity is not necessarily confined to the connections between person-events or temporal person-stages, but also covers links between people. Chapter IV effectively claims that integrity can even be extended from simple, two-person examples to more complex associations like the state. So the case for corporate personality now stands or falls with an argument showing that this idea has practical purchase where ethical individualism falls short. I look for such an argument in Chapters V, VI and VII.

CHAPTER V

EXPANDING OUR ETHICAL HORIZONS

In order that there may be institutions, there must be a kind of will, instinct, or imperative, which is anti-liberal to the point of malice: the will to tradition, to authority, to responsibility for centuries to come, to the solidarity of chains of generations, forward and backward *ad infinitum*. When this will is present, something like the *imperium Romanum* is founded [...] The whole of the West no longer possesses the instincts out of which institutions grow, out of which a *future* grows: perhaps nothing antagonizes its 'modern spirit' so much. One lives for the day, one lives very fast, one lives very irresponsibly: precisely this is called 'freedom'.

Friedrich Nietzsche – *Twilight of the Idols*¹

Introduction

The argument for corporate personality has two basic parts. The first part of the argument maintains that the concept of integrity fits groups as much as it fits individuals. Integrity does not only explain how different person-stages should be connected to form one person over time. It also explains how different persons should be connected to form one group person. What connects person-stages in the relevant way is the fact that their present aims are guided by future-directed intention. What connects persons in the relevant way is the fact that their individual aims are guided by a collective intention.

The second part is normative. It asserts that integrity is a virtue as much for associations as for individuals. It is important to unpack this claim a bit further. In Chapter II, I followed Rawls, Korsgaard and others in arguing that the concept of personal identity should be understood as a regulative ideal that is immanent in a wider normative practice or tradition, which is itself not again grounded in a metaphysical

¹ F. Nietzsche, *Twilight of the Idols: Or How One Philosophizes With a Hammer*, in: W. Kaufmann (ed.), *The Portable Nietzsche* (Harmondsworth: Penquin, 1984), p. 543.

theory of the self. In other words, our understandings of personal identity should be judged on the basis of reflective equilibrium, like normative theory in general. Now, the practice theory stands ethical individualism in good stead in holding off Parfit's impersonal challenge. It points to an irreducibly normative argument for the moderate claim. So even though personal identity is relational, we still have good reasons to claim that personal identity has a basic normative input. However, the problem for individualism, I now want to argue, is that the practice theory could topple the argument over to the other side. We already know that it is possible to claim that associations are composed of a kind of normative relatedness that is comparable to personal integrity. I earlier referred to this claim as *associativism*. So if the practice theory is plausible, then we may also be able to construct corporate personality as a regulative ideal immanent in a practice or tradition. In other words, the idea of a corporate person could still prove its credentials as a normative conception in reflective equilibrium with our basic intuitions and well-considered judgements.

In Chapters V, VI and VII, I want to argue that corporate personality can indeed be seen in the same, normative vein as personal identity. This argument seeks to show that there are certain practices and types of aspiration, which also form part of our form of life, our way of being, that *presuppose* a sense of narrative unity and continuity that only associations could provide. These practices and aspirations cannot be fully captured by ethical individualism, the view that only individual goods and preferences enjoy moral and rational standing as reasons. They presuppose that we conceive not only of individuals, but also of associations or groups as the *loci* of independent interests and value, as lives with a unity and purpose, in sum, as corporate persons.

What does this argument look like? Let me briefly sketch the outlines of my answer. I begin by considering two familiar hypotheses. Section 1 briefly considers the notion that cultural goods are 'irreducibly social'. Section 2 submits the view that associativism provides us with a proper account of human sociality or co-operativeness.

Both ideas face important objections. Section 3 then switches to another, though clearly related idea. I want to argue that corporate personality lengthens the temporal horizons of agency in a way that ethical individualism is incapable of. The idea of corporate personality has the benefit that it gives us a theoretical foothold for explaining how past, present and future generations are interconnected in a normative sense, or so I shall argue.

This hypothesis has a negative and a positive component. Sections 4 and 5 elaborate on the negative component of the argument by focussing on the issue of intergenerational justice. Theories of distributive justice generally aspire to the conclusion that scarce resources ought to be distributed fairly over generations. The challenge, which they therefore set themselves, is to show that present generations have a duty of justice to pass on resources to future generations. I argue that the two main strands of this justice-based argument fail to meet this challenge.

Chapter VI then explores the positive component. It demonstrates that corporate personality evades the difficulties that haunt the individualist view of intergenerational justice. Finally, Chapter VII investigates how the argument could be extended to another transgenerational issue, namely how past generations can bind present and future generations by means of constitutions and agreements to third parties.

1. Irreducibly Social Goods

I want to begin by mentioning what might seem the most evident argument for corporate personality and is commonly subsumed under the banner of communitarianism. Briefly, the communitarian argument, as initiated by MacIntyre, Taylor and Sandel in the nineteen-eighties, alleges that political theory has succumbed to a fixation on individual rights, goods and interests. This, moreover, means that most theorists have ignored the vital importance of culture, language and national identity.

Communitarians argue that such ties and associations are in more than one way essential to the flourishing of human life. In fact, the communitarian critique is supported by two closely related, but separate pillars. The first pillar establishes that culture and language are an essential ingredient to human well-being. Here, the communitarian thesis points out that the human capacity for making *meaningful* choices presupposes a context of background meanings and understandings provided by language, history, tradition and culture.² The second pillar argues that such ‘communal goods’ are essentially and ‘irreducibly social’ and cannot be accommodated within the theoretical framework of ethical individualism. This last claim, holds that the background goods of language and culture are *irreducibly social or participatory*.³

Here, I do not have the space to discuss the issue at length, but as it stands, the communitarian thesis remains debatable. The controversy is not so much about the cogency of the first part of the argument. What is problematic, instead, is that individualism seems equally capable of articulating the importance of background goods. In recent years, it has become increasingly evident that so-called ‘liberal communitarian’ approaches have successfully reconciled a concern for culture with the individualist foundations of liberalism. For example, Buchanan has argued that individual rights to association, life, property, religion, thought and speech form a sufficient guarantee for civil society and communal life to flourish.⁴ More notably,

² See for example C. Taylor, ‘Lichtung or Lebensform’, pp. 61-78; ‘The Importance of Herder’, pp. 79-99, both in: *Philosophical Arguments*.

³ See especially C. Taylor, ‘Irreducibly Social Goods’, pp. 127-145; ‘Cross-Purposes: The Liberal-Communitarian Debate’, pp. 181-203, both in: *Philosophical Arguments*. The same arguments are made by: J. Waldron, ‘Communal Goods as Human Rights’, in: *Liberal Rights* (Cambridge: Cambridge University Press, 1993) pp. 339-369; L. C. Green, *The Authority of the State* (Oxford: Clarendon Press, 1988), pp. 206-209; D. Réaume, ‘The Group Right to Linguistic Security: Whose Right, What Duties’, in: J. Baker, *Group Rights* (Toronto: University of Toronto Press, 1994), pp. 118-141; A. Mason, *Community, Solidarity and Belonging* (Cambridge: Cambridge University Press, 2000), pp. 43-45.

⁴ A. E. Buchanan, ‘Assessing the Communitarian Critique of Liberalism’, *Ethics* 99/1989, pp. 852-882; ‘Liberalism and Group Rights’ in: J. L. Coleman & A. E. Buchanan (eds.), *In Harm’s Way: Essays in Honor of Joel Feinberg* (Cambridge: Cambridge University Press, 1994), pp. 1-15. For a related approach, see C. Kukathas, ‘Are There Any Cultural Rights’, *Political Theory*, Volume 20 (1992), pp. 105-139.

Kymlicka offers an argument for background goods from a liberal-egalitarian angle.⁵ In his view, culture and language are on a level pegging with conventional primary goods or resources or capabilities. Like Taylor, Kymlicka is keenly aware of the significance of culture. However, it is precisely for the reason that cultural variables affect individual autonomy, that justice demands that we look, not only at the distribution of material resources, but also of cultural resources. Both Buchanan and Kymlicka point out that individuals have stakes in background goods, and they incorporate a concern for these goods in the framework of liberalism. To put it differently, the key problem is that the communitarian argument from the good of culture, language and belonging collapses into ethical individualism.

2. Prudence and Beyond: Extending the Scope of Reason

I want to consider another well-known hypothesis, which may be linked to Hume, Burke and Hegel.⁶ The general idea is that our involvement in associations and groups is linked to why people act morally in the first place. Here, I use the term moral in the very general (and non-Kantian) sense of sociable, co-operative, civil and reciprocally responsive to other people. Burke once put the idea as follows:

To be attached to the subdivision, to love the little platoon we belong to in society, is the first principle (the germ as it were) of public affections. It is the first link in the series by which we proceed toward a love to our country and to mankind.⁷

⁵ W. Kymlicka, *Liberalism, Community and Culture* (Oxford: Clarendon Press, 1989). See also W. Kymlicka, 'Individual and Community Rights'.

⁶ See also R. Sugden, 'Beyond Sympathy and Empathy: Adam Smith's Concept of Fellow-Feeling', *Economics and Philosophy*, 18 (2002), pp. 63-87; M. Hollis, *Trust within Reason* (Cambridge: Cambridge University Press, 1998); K. Graham, 'Morality, individuals and Collectives'; D. Miller, *On Nationality*, pp. 57-58.

⁷ E. Burke, *The Reflections on the Revolution in France*, J. G. A. Pocock (ed.), (Indianapolis: Hackett Publishing Company, 1987), p. 41.

Briefly the hypothesis claims that the reason why we extend our concern from a narrow focus on our present aims, to a wider view of our temporally extended good, and then to an even wider concern for the interests of other people, derives from the fact that, in various ways and to various degrees, we participate in associations like the family and the nation-state. On this view, then, the possibility of a *corporate* person emerges because we have reason to treat the interests of our fellow-members as fundamental, just as the idea of an *individual* person emerges because you have reason to treat the interests of your future selves as fundamental.

Consider again what the standard model of prudential reason claims. This model argues that what an agent has reason to do derives from the interests of both his present and future selves. That is, we have reason to do whatever promotes the goodness of our lives conceived as temporal wholes. Obviously, we may still show appropriate concern for the interests of other people, but this concern must be justifiable in terms of our own good. The essence of the model is that it leaves no room within the scope of reason for the non-derivative importance of the good and preferences of other people. The associative view, in contrast, claims that there is such room – albeit only to the extent to which an agent is connected with these other people.⁸ But while the standard model of prudence conceives of reason more narrowly and associativism more broadly, both positions spring from a common source.

⁸ This claim, then, falls well short of the stronger, Kantian view, which argues that there is room even for the non-derivative importance of others *per se*. Nagel, in his book *The Possibility of Altruism*, argues his case for the latter view by analogy to the *standard* view of prudence. This account of prudence, Nagel says on page 87, ‘derives from the requirement that reasons be timelessly formulable’. He believes (p. 88) that morality is ‘closely analogous to the formal principle of timelessness, in that it will deny the possibility of restricting to one *person* the derivative influence of a reason for action, just as the formal principle which underlies prudence denies the possibility of restricting such derivative influence to one *time*.’ While I am arguing for something like ‘the possibility of corporate personality’ on the basis of a similar comparison, the difference between Nagel’s view and mine is that I reject his idea of objectivity, *viz.* of timelessness and impartiality. Nagel thinks that since we should conceive of prudential reason as timeless, it is equally possible to conceive of reasons as agent-neutral. I do not accept the premise of that argument. My claim is that since prudence awards reason-giving value to *intrapersonal* connectedness through time, we can also conceive of an argument that assigns such value to *interpersonal* connectedness. However, the reasons in both cases are not ‘objective’ in Nagel’s sense as they clearly depend on the relational fact of personal identity.

Recall what Parfit claims. If the Humean view of personal identity is correct, then we could conceive of an argument that pushes prudential reason one step further back. Parfit's present-aim theory suggests that we have reason to do, not what achieves the good of a temporally extended self, but whatever realises the aim or preference of our current person-stage.⁹ If this theory were indeed correct, then the scope of reason would have to be circumscribed still more narrowly. So, on the one hand, we have the present-aim theory, which leaves no room for the non-derivative importance of the interests or preferences of our future selves. On the other hand, we have the orthodox view of prudence, which claims that there is such room. This view claims that the preferences or interests of our future selves count as reasons in our present deliberations.

In claiming this, the standard view of prudence turns out to be merely a special brand of associativism. Given the discussion in Chapter II, this 'micro-associativist' view is easily clarified. On the Humean theory of personal identity, the self essentially *is* a temporal association. Why still care about my future well-being or preferences? Well, I have reason to care about my future because (1) I still form a psychological bundle with future person-stages $S_{t+1}, S_{t+2} \dots S_{t+n}$ as set out above, and (2) this relational fact matters in inherently reason-giving ways. Since we may now have to allow a discounting rate in proportion to the strength of the relatedness, the result is not quite the same as on the standard account of prudence. But I do have reason to care.

If the combination of (1) and (2) extends the scope of reason to *prudential* concern, then associativism extends the scope of reason to *moral* or other-regarding concern. Why care about the well-being or preferences of others? According to associativism, the reason stems from (3) the fact of our psychological connectedness with these others, and (4) the claim that these relational facts matter in a similarly reason-giving manner. So I have reason to care because the well-being of *my* fellow-associates is at stake. This interpersonal connectedness gives me a reason to give to others the kind of non-

⁹ The present-aim theory is set out and defended in *Reasons and Persons*, Part 1 and 2. For the identity-based version of the argument, see especially Chapter 14.

instrumental respect that the prudential model awards to my future selves. To put it differently still, associativism claims that the preferences and interests of connected others are directly reason-giving for myself.

The gist, then, of the associativist thesis is that our normative horizons are extended from the absolute minimum of present-aim theory, first to prudence, and then from prudence to morality, by the same underlying idea, namely that psychological relatedness, whether *intra* or *interpersonal*, matters in reason-giving ways. The interests, aims and preferences of future selves constitute reasons for us now because these future selves are in some way connected to us. Likewise, claims (3) and (4) explain our triumph over narrow selfishness. If this understanding of interpersonal relatedness is correct, the divide between self and other becomes less sharp, not only in the metaphysical but also in the normative sense. The gap between self and other-regarding reasons is bridged by the claim that the interests of the one count as reasons for the other in much the same way as future interests count as reasons now.¹⁰ This effect is particularly strong in the case of family ties and friendship. But as the circle is drawn wider, as we include our country and even humankind, it gradually loses its force in proportion to the weakening strength of the connectedness. Another way of putting it is to say that associativism involves a social discounting mechanism. We do not have to give the interest of everyone an equal weight, but are allowed allocate our resources roughly in accordance with the strength of our connectedness to others.¹¹

How does this relate to the idea of corporate personality? As I have been at pains to show, we should accept or reject conceptions of personal identity on the basis of

¹⁰ David Brink's conception of 'metaphysical egoism' advances precisely this idea. See his 'Self-love and Altruism', *Social Philosophy and Policy*, Volume 14, Number 1 (1997), pp. 122-157; 'Rational Egoism and the Separateness of Persons', in: J. Dancy (ed.) *Reading Parfit*, pp. 96-134.

¹¹ Hume argued 'that where, besides the general resemblance of our natures, there is any peculiar similarity in our manners, or character, or country, or language, it facilitates the sympathy. The stronger the relation is betwixt ourselves and any object, the more easily does the imagination make the transition, and convey to the related idea the vivacity of conception, with which we always form the idea of our own person.' In *A Treatise of Human Nature*, p. 261. I return to the issue of the discounting mechanism in Chapter VI.

reflective equilibrium. Part of the appeal of ethical individualism, I believe, is that humans tend to be deeply concerned about their future. We could, on the practice theory, still set aside the notion of prudential reason. But since this concern is so crucial to *our* form of life, it seems more plausible to embrace a view of the person that is tailored to fit this practice. The idea of integrity is in the end simply more persuasive than the notion that the temporal parts of a person are distinct agents. Could it be argued that the idea of a corporate person has normative appeal because it extends the scope of reason to other-regarding accounts of reason in the way just proposed? After all, we know that human beings are capable of co-operation and compassion and generosity. And we know that our relatedness to others plays a significant part in this. As Hume remarked,

... it appears, that in the original frame of our mind, our strongest attention is confin'd to ourselves; our next is extended to our relations and acquaintance; and 'tis only the weakest which reaches to strangers and indifferent persons. This partiality, then, and unequal affection, must not only have an influence on our behaviour and conduct in society, but even on our ideas of vice and virtue; so as to make us regard any remarkable transgression of such a degree of partiality, either by too great an enlargement, or contradiction of the affections, as vicious and immoral.¹²

Associativism, then, seems as deeply embedded in our considered judgements and intuitions as our preoccupation with our selves. What is more, it has an important impact on our ethics.

That said, this hypothesis faces at least two familiar objections, which are both grounded in versions of what I shall refer to as contractualism. The first objection holds that associative relatedness is valued too highly. Briefly, it is claimed that things like social co-operation, justice, trust and reciprocity could well derive their justification from the prudential interests and preferences of individuals alone.¹³ Assuming this to be

¹² D. Hume, *A Treatise of Human Nature*, p. 488.

¹³ What follows is of course a very coarse synopsis of what is an extremely rich and powerful idea going back at least to Hobbes's *Leviathan*, Tuck, R., (ed.) (Cambridge: Cambridge University Press, 1991). Influential contemporary contributions to this tradition in political theory include: J. M. Buchanan, *The Limits of Liberty: Between Anarchy and Leviathan* (Chicago: University of Chicago Press: 1975); D.

the case, the idea that wider associations and interpersonal relatedness have primitive, reason-giving weight is simply unwanted theoretical surplus. Norms of civility and justice, while ostensibly unselfish in intent, are in fact merely agreements or conventions serving the self-regarding preferences of all the participants involved. We could call this the 'mutual advantage objection'. Essentially, this objection suggests that we can explain reciprocal or socially responsive behaviour in terms of *agent-relative* justifications. Such justifications are rooted in the agent's preferences and desires.

Notice that this agent-relative justification is more plausible if the agent's preferences and interests are defined over the *long term* rather than the *short-term*. If we assume that the agent is only disposed to his present interest or aims, it becomes more difficult to find a satisfactory answer to Hobbes's 'Foole'. That is, it becomes more difficult to see how rational agents could benefit from behaving in a co-operative and civil manner.¹⁴ Thus, this justification takes for granted that an agent is predisposed to the interests and aims of his future selves. The difference with the wider associativist view above is merely that this perspective is confined to the relatedness between person-stages. While it is affirmed that relatedness over time has reason-giving weight, it is denied that the relatedness between agents matters similarly.

The second contractualist objection moves to the level of *agent-neutral* justification. I shall refer to it as the Kantian objection.¹⁵ Essentially, it points out that we should carefully distinguish between, on the one hand, the problem of compliance with the

Gauthier, *Morals by Agreement* (Oxford: Oxford University Press, 1986); G. S. Kavka, *Hobbesian Moral and Political Theory* (Princeton: Princeton University Press, 1986); J. Hampton, *Hobbes and the Social Contract Tradition* (Cambridge: Cambridge University Press, 1992).

¹⁴ The problem of the 'Foole' is famously captured by the Prisoner's Dilemma. A common suggestion is that it becomes less pressing once a long-term perspective is introduced by iteration of the game. The *locus classicus* is R. Axelrod, *The Evolution of Social Co-operation* (New York: Basic Books, 1984). Alternatively, it has been claimed that rational agents have reason to conditionally commit themselves, in the manner of the planning model of agency, to acting co-operatively in the future, and hold to this commitment even when this would be contrary to their present interests and aims. See E. McClennen, 'Rationality and Rules', in: P. A. Danielson (ed.) *Modelling Rationality, Morality, and Evolution* (Oxford: Oxford University Press, 1998), pp. 13-40.

¹⁵ See here for example T. Nagel, *The Possibility of Altruism*; B. Barry *Theories of Justice* (London: Harvester Wheatsheaf, 1989).

standard and, on the other, what the standard of *pure* reason entails as such. Clearly, at the compliance level, people are primarily motivated to concern themselves with the well-being of relatives and fellow-members. But the Kantian objection plays down the significance of this fact to theories that aim to explain what we ought to do. Fundamentally, it is claimed, we *have* no reason to attach more importance to our future selves or our relatives and fellow-countrymen than to perfect strangers. Reason ultimately dictates that we concern ourselves with the interests of other people, not because we are fellow-associates, but because each individual agent forms an end in and of itself. At the deepest, normative level, the structure of reason is always objective and neutral. My own interests and preferences produce no greater reason than those of any other individual person. Of course, it may be true that people do not naturally comply with the demands of pure impartiality, but the principle is no less valid. Reason demands that all individuals weigh equally in our judgements, whether this principle corresponds to our *de facto* motivations or not. To put it differently, while we may live in a non-ideal world, where we may find it hard to act in full accordance with reason, this does not show that this view falls short at the ideal level.

Given the considerable weight and pedigree of these objections, it is less than straightforward to find conclusive support for associativism. Far be it for me to argue that contractualism is triumphant, but it seems to me a fair criticism that in view of these alternatives the idea of corporate personality should be able to clearly demonstrate its superiority as a normative political conception. The notion has to make plain why its account of the civil order is superior to its rivals.

Thus far, the concept of a collective intention or preference, which I discussed in Chapter III, has been employed mainly to show how social co-operation and co-ordination come about in general. For example, Hollis, in his book *Trust within Reason*, has argued that, under certain conditions, both versions of the contractual model fail to provide a plausible account of the social order. According to Hollis, society would

implode without the prevalence of a collective attitude that he describes as *trust*. On the one hand, the motivational assumptions of game theory, which prevent genuine trust from emerging, can be seen to encourage sub-optimal results in a spate of circumstances.¹⁶ Full-blown Kantianism, on the other hand, while avoiding some of these obstacles, does not capture the conditional and also personal nature that is typical of trust. And so associativism emerges as a third and winning theory of reasonableness.

Although I am sympathetic to Hollis's view, I nevertheless suspect that ethical individualists would wish to express serious reservations with the gist of his arguments. Doubts linger as to whether the mutual advantage model really falls short of providing an explanation for the civil order at this level of generality. As is well known, the model is often claimed to depend on the introduction of a long-term perspective, but Hollis rejects these claims perhaps too easily as mere 'technicalities'.¹⁷ By the same token, it is not all that evident that Kantian impartiality does indeed fail to grasp the personal nature of trust and human relationships. According to one common view, at least, impartiality should be understood to operate, not on the day-to-day level of human affairs, but only at the higher level of justification.¹⁸

¹⁶ The problem with rational choice, Hollis writes on page 68, is that agents 'are like two methodological solipsists or windowless monads, each with an individual point of view containing the other's within it. There is no corporate body whose members they are. When Adam is trying to thread a needle, success does not depend on separate strategic choices by his right and left hands, each deciding where to best position itself, after considering what the other will do, since all choices are Adam's. But this offers no analogy for harvesting as a corporate action, since Adam and Eve are not a corporation [...]'. In the same spirit, Robert Sugden appeals to 'team-reasoning' to solve co-ordination games, which standard game theory is said to leave unresolved. The problem with co-ordination, says Sugden, is that rational agents have to second-guess each other's choices, leading to infinite regress. My reasoning depends on your reasoning, which in turn depends on my reasoning, and so the cycle continues. The key advantage of team-reasoning is that it overcomes this problem. According to Sugden, team players deliberate like a corporate agent, with just one collective preference or utility function to guide their reasoning. See his articles 'Team Preferences', pp. 181, 191; 'Thinking as a Team'; 'Rational Co-ordination', in: F. Farina, F. Hahn, and S. Vannucci (eds.) *Ethics, Rationality, and Economic Behaviour*, pp. 244-262. Similarly, but now in the case of the one shot Prisoner's Dilemma, Hurley argues, in her *Natural Reasons*, Chapter 8, that agents will co-operate only if they assume the perspective of collective agency. See also M. Hollis and R. Sugden, 'Rationality in Action', *Mind*, Volume 102, Issue 405 (January 1993), pp. 1-35; M. Bacharach, 'Interactive Team Reasoning: A Contribution to the Theory of Co-operation', *Research in Economics*, Volume 53, No. 2 (June 1999), pp. 117-147; R. Tuomela, *Cooperation: A Philosophical Study* (Dordrecht: Kluwer Academic Publishers, 2000).

¹⁷ M. Hollis, *Trust within Reason*, p. 59.

¹⁸ *Ibid.*, Chapter 5. See my discussion of two-level impartiality in Chapter VII.

Now, we could decide to dig in our heels at this point and hope to be able to demonstrate that these doubts are in the end unfounded. However, in the remainder of this thesis, I would like to explore a more specific variation on this theme instead. Albeit at a lower level of generality, this variation provides us with a more clear-cut case for associativism, or so I argue. It points out that associations tend to have a greater continuity than individual persons. Due to this greater continuity, associations are not merely essential to *broaden*, but also, and perhaps mainly, to *lengthen* our practical horizons in time. Associations do not merely extend our concerns and pursuits to our contemporaries, but also to the past and future.

3. Conservatism

I have already mentioned four different theories of the scope of reason. To begin with, Parfit's present-aim theory confines reason to our own present interests or preferences. Second, the standard account of prudence extends the scope of reason to our own future interests or preferences. Third, associativism widens that scope to include the interests and preferences of others, but mediated by our social connectedness. Fourth, moral impartiality argues that, at the deepest level, the good and preferences of all individual agents weigh equally as reasons. We can now add a fifth view. This is that associativism allows for an extension of that scope to the interests and preferences of past and future generations.

On a historical note, the idea that associations make possible such a trans-generational extension is closely affiliated with Burke's conservatism.¹⁹ In *The Reflections of the Revolution in France*, Burke advanced a conception of the state as an interconnecting chain between generations past, present and future. One of his objections to the French Revolution was precisely that by breaking free from this chain,

¹⁹ See also R. Scruton, 'Corporate Persons'.

all that previous generations had achieved and accumulated would stand to return to 'the dust and powder of individuality' and disperse 'to all the winds of heaven'. As Burke wrote: 'The perennial existence of bodies corporate and their fortunes are things particularly suited to a man who has long views; who mediates designs that require time in fashioning, and which propose duration when they are accomplished.'²⁰

Now, recall that we rejected the extreme claim because we all have and need certain ground-projects and conceptions of the good, and because we are all inevitably immersed in relationships, which presuppose rights, obligations and responsibilities with temporal continuity. And so we ended up supporting the moderate claim. Here, I want to argue that the notion of corporate personality owes part of its appeal as a normative conception to the fact that we have even longer lasting projects and relationships, namely those that involve the participation of several generations. It is important to present this argument in the correct order because it comprises two separate parts. The first part is I think relatively uncontroversial and largely at the pre-theoretical level of intuition. It maintains that co-operation and relations between generations are part and parcel of the same practices and goings-on to which the ethical individualist has to appeal in order to defeat the extreme claim. Most people do not just care about themselves or about other people in their vicinity, they also look back and forward in time, and much further than their own lives. We find ourselves not merely absorbed by our own private plans and projects, but also by bigger plans and projects with their own history and continuity.

This is I think unquestionably true in the case of the past. In his wide-ranging study of our attitudes towards the past, Lowenthal writes:

The miracle of life is cruelly circumscribed by birth and death; of the immensity of time before and after our own lives we experience nothing. Past and future are alike inaccessible. But, though beyond physical reach, they are

²⁰ E. Burke, *The Reflections on the Revolution in France*, p. 139.

integral to our imaginations. Reminiscence and expectation suffuse every present moment.²¹

Elsewhere he points out: ‘Popular history, biography, autobiography, and historical fiction deluge bookshops and television screens; classroom re-enactments emphasize the immediacy and malleability of past events; ‘family-history’ has transformed genealogy from an elite pursuit into a popular preoccupation; oral archives celebrate the annals of hitherto unsung masses and local histories their most humdrum details’.²² The streets of Western capitals are generally littered with monuments commemorating past figures and events. Lord Nelson still stands high on his column in Trafalgar Square. Together with countless others, his statue is the material reminder of the achievements of and the tragedies suffered by our ancestors.

More importantly, the past is not merely of academic and aesthetic interest. We attach practical and political consequences to history. Sometimes we feel an enormous sense of pride or indignation in connection to events that took place long before we were born. Examples abound: We honour those who gave their lives in the World Wars; the German state still accepts its war debt, while certain German corporations do the same; people demand apologies and reparations over alleged crimes committed by colonial powers and slave-trading organisations centuries ago; Greece is still adamant that the Elgin Marbles be returned from their current location in the British Museum in London; nationalist movements use and abuse past events for inspiration and justification of their present ends; past generations pass on their national debts to future generations; and century-old constitutions, laws and treaties continue to bind us today. I am not contending, of course, that all these claims and practices are perfectly good and reasonable. What I am claiming, however, is that we commonly affirm that our relatedness to the past is generally not unlike our relatedness to an earlier self. The past, indeed, is an integral part of us. We do not just draw a line under history, but consider

²¹ D. Lowenthal, *The Past is a Foreign Country* (Cambridge: Cambridge University Press, 1985), p. 3.

²² *Ibid.*, p. 367.

ourselves bound to our ancestors. Our attitudes to the past are very complex and have themselves undergone many changes. Lowenthal suggests that to an important extent these complexities and changes pertain to how we have come to evaluate the past. But whether we characterise the past as a benefit or a burden, whether we exhort the past as an example or reject its authority, we are historical creatures who tend to frame the issues that face us today within the context of a conception of time that exceeds our own brief lives and connects us with earlier generations. History may sometimes stand in the way of our happiness, but the idea that generations are unconnected – like ‘the flies of a summer’, as Burke put it – does not appeal to us.

The past is not a foreign country but ‘assimilated in ourselves’, Lowenthal quite rightly concludes.²³ What is less obvious, perhaps, is that this assimilation also extends itself in a forward direction. By linking ourselves with the past, we also link ourselves with the future. Burke saw that both perspectives spring from the same source. ‘People will not look forward to posterity,’ he said, ‘who never look backward to their ancestors.’²⁴ Both perspectives appear as two sides of the same coin. Our connectedness with the future is perhaps less tangibly present than our assimilation with the past. Nevertheless, it still has an important impact on our practical deliberations. Many people take an interest in future states of the world, even though they are certain never to witness these states themselves. Most of us, I believe, are not wholly indifferent to what happens after our deaths. For example, we express concern about climate changes in the next hundred years. Of course, some people allege that such concerns are not justified. But typically these assertions are grounded in the belief that the effects of climate change will be less harmful than some fear. There are very few people – let alone policy-makers – who would openly argue that climate change constitutes no reason for concern simply because the adverse effects thereof will be noticeable only

²³ Ibid., p. 412.

²⁴ E. Burke, *The Reflections on the Revolution in France*, p. 29. The idea that a forward looking perspective goes naturally together with a backward looking perspective is affirmed in Bratman’s planning model of agency, which I discussed in Chapter II.

after our deaths.²⁵ In a similar vein, people and governments are concerned to conserve buildings, ancient cities and valuable landmarks for the future.

Elster claims that human reason is comparable to that of a 'globally maximizing machine'.²⁶ Other animals are 'gradient climbers', who settle for the first 'local equilibrium' that they come across. But humans investigate the possibility of more rewarding equilibria in the future. So in contrast to other organisms, humans think several steps ahead. Moreover, they have acquired the crucial capacity to reject instantaneous gratification in order to reap a greater gain. We can see the point of taking one step backward in order to take two steps forward, even though we may have to face up to the problem of weakness of will. Now, according to Elster, this perspective of the future is constitutive of our very idea of personhood. He is probably right. But my point is that human reason is not confined to a global equilibrium within one life or one generation. There is a sense in which even the future beyond our own lives matters to us. And this last perspective, I want to suggest, is no less constitutive of our idea of personhood.

Here, it is important to reiterate that I do not intend to ground the claim that we *ought* to care about the future in an antecedent view of corporate personality. I merely want to point out that a sense of continuity between past, present and future is implicit in many of our political and moral practices. And the fact that we have such judgements and intuitions underpins the idea of corporate personality as a normative conception. Thus the case for associativism, like the case for the moderate claim, emerges only when we begin from a normative standpoint and we accept certain pre-theoretical judgements from the outset.

²⁵ At the so-called 'Earth Summit' in Rio de Janeiro the United Nation reached agreement on the principle that 'The right to development must be fulfilled so as to equitably meet developmental and environmental needs of present and future generations.' See the UN *Report of the United Nations Conference on Environment and Development*, at <http://www.un.org/documents/ga/conf151> (1992).

²⁶ J. Elster, *Ulysses and the Sirens*, Chapter I.

In what follows, I shall assume that the idea that future generations matter is relatively uncontroversial at the intuitive level. Conservatives, insofar as they are worthy of their name, would agree. More importantly, it is also evident that individualist theorists like Rawls, Barry, Kavka and Gauthier tend to share this intuition. Hence, I want to carry on with the second part of my argument, which is predominantly theoretical. Briefly, this maintains that the idea of corporate personality is necessary to corroborate this intuition within a plausible theoretical framework. In order to make this claim stick, it must first be argued that the mutual advantage and Kantian objections to associativism lose their gravity in the trans-generational context. I want to press home the negative part of this argument in the remainder of this chapter, mainly by focussing on the idea of justice between generations.

4. Reciprocity and Obligations to Future Generations

As I just said, I want to concentrate on the claims that generations could put on each other in terms of justice. In recent decades, such claims have been widely discussed, mainly in relation to the depletion of resources, environmental pollution, climate change, the accumulation of national debt, and so forth. Similar claims, I think, may be advanced for the conservation of cultural resources. But to keep matters simple, I shall build on an example of Parfit's.

The Depletion Case: The present Generation *A* discovers and controls a non-renewable but highly valuable resource *X*. There are two options: Generation *A* could nearly exhaust *X*. Near depletion maximises utility for *A*, it leaves Generation *B* worse off, and it means that Generation *C* will eventually face a catastrophe. Alternatively, *A* could conserve sufficient *X*, allowing *B* and *C* to co-profit from the resource. Conservation means that both *A*, *B* as well as *C* have an equal share of this resource. Consequently, while *A* does less well than it would have done otherwise, *B* and *C* benefit considerably.

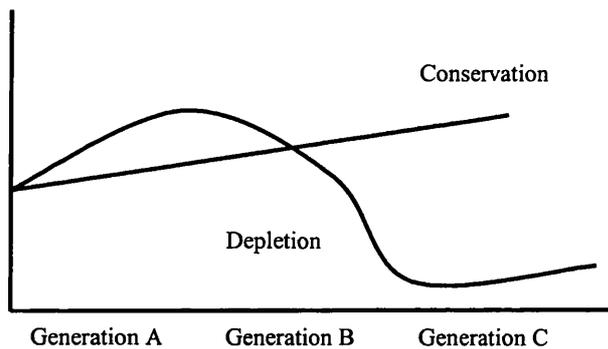


Figure 1 *The Depletion Example*. Source: D. Parfit, *Reasons and Persons*, p. 362

Now, I shall assume that most people agree intuitively that Generation *A* ought to weigh the interests of *B* and *C* in its decision. These people believe that *A* ought to decide for conservation. But the question is how to validate this intuition within the framework of ethical individualism. And for this, it is natural to turn to the idea of distributive justice.

To begin with, consider the argument from justice as mutual advantage, which essentially claims that conservation provides a Pareto-improvement leading to gains for all. On first inspection, this seems clearly false. While it leads to gains for *B* and *C*, and may even maximise overall utility, conservation leads to a loss for Generation *A*. This problem arises partly because *A*, *B* and *C* find themselves in conditions devoid of reciprocity and equality. As has often been noted, there is nothing that a future generation could do for us now.²⁷ Consequently, conservation imposes a costs on *A* for which it could receive no compensation. As Heilbroner once put the problem: ‘Why should I lift a finger to affect events that will have no more meaning for me seventy-five years after my death than those that happened seventy-five years before I was born?’²⁸ He finds it very hard to imagine a ‘rational’ answer in this case.²⁹

²⁷ A democratic process of preference aggregation equally fails to capture future interests. See J. G. March and J. P. Olsen, *Rediscovering Institutions: The Organizational Basis of Politics* (New York: The Free Press, 1989), pp. 146-149.

²⁸ R. L. Heilbroner, ‘What Has Posterity Ever Done For Me?’, in: E. Partridge (ed.) *Responsibilities to Future Generations: Environmental Ethics* (Buffalo: Prometheus Books, 1981), pp. 191-194, at p. 191.

²⁹ For a frank assessment of the limitations of contractarianism in this context, see G. S. Kavka, *Hobbesian Moral and Political Theory*, especially pp. 443-446. For a further critical analysis, see also D. Braybrooke, ‘The Social Contract and Property Rights across the Generations’, in: P. Laslett and J. S.

4.1 Generational Overlap

However, in his *Morals by Agreement*, Gauthier sets out to solve the problem by pointing out that generations overlap: 'Each person interacts with others both older and younger than himself, and enters thereby into a continuous thread of interaction extending from the most remote human past to the farthest future of our kind.'³⁰ Let us say, then, that Generation *A* overlaps with *B*, and *B* with *C*, and so forth. Let us further grant, as seems plausible, that since their lives are partially overlapping, *A* and *B* could agree on a fair distribution of *X*. For example, *B* would get a share if it provides for a pension scheme in return. Now, while there is no direct reciprocity between *A* and *C*, the key to Gauthier's solution is that this reciprocity will be mediated by *B*. Thus, *C* can benefit or harm *B*, and *B* can benefit or harm *A*, and so they all connect. It seems that he believes that the interests of posterity can be adequately secured in this way. *B* knows that it could benefit from the same agreement with *C*. So, as *B* negotiates terms with *A*, it 'must keep in mind [its] need to establish similar terms with those of a later generation'.³¹

Notice that Gauthier claims that the terms of the contract must necessarily be *similar* or the *same*. In order to be able to gain from an agreement with *C* later, *B* has an incentive to insist on terms similarly acceptable to *C*, and it would further know that *C* would insist on terms similarly acceptable to Generation *D*, and so forth. Hence, some equitable policy would result. Or so at least the theory claims. For example, *A* might initially suggest to *B* that they divide the resource between them. But *B* could not accept this proposal because it knows that it will eventually need to secure co-operation with *C*. *A* then proposes a three way split. But still, *B* also knows that *C* will eventually need

Fishkin (eds.), *Justice Between Age Groups and Generations* (New Haven: Yale University Press, 1992), pp. 107-126.

³⁰ D. Gauthier, *Morals by Agreement*, p. 299.

³¹ *Ibid.*, p. 299, emphasis added. For a similar view, see J. Heath, 'Intergenerational Cooperation and Distributive Justice', *Canadian Journal of Philosophy*, Volume 27, Number 3 (1997), pp. 361-376.

to co-operate with *D*, and *D* with *E*. Consequently, we arrive at an equal cut for *all* future generations. If there are an *n* number of generations, each generation obtains $1/n^{\text{th}}$ share.

However, while intriguing, this solution is in the end not very credible. On closer inspection, there appears to be no persuasive reason why *B* must offer the exact same terms to *C*. Why not simply negotiate one deal with *A* and reach an entirely different agreement with *C*, with the aim of profiting optimally from co-operation with *A* and *C* respectively?

Assume again that overlapping generations *A* and *B* are trying to find a scheme that is mutually beneficial to them both. While *A* fully controls access to resource *X*, *B* could offer certain benefits in return for a share. Now, surely it is not beyond the realm of possibilities that *A* and *B* agree to use, say, 40 percent each and leave only 20 percent to future generations. In fact, such an agreement is *prima facie* plausible because *A* would still control a good part of *X*, while gaining from *B*'s co-operation in return for saving the remainder. *B* still receives a far greater share than it would have received either if *A* had taken it all or the resource would be shared equally between all future generations. Moreover, *B* would still be able to profit from a degree of co-operation with *C*. When it arrives on the scene, *B* could tell *C*: '*X* is now almost exhausted and we have benefited nicely from our contract with *A*. If you want what is left, then you would have to do this for us in return.' The terms of this agreement are not the same as the one between *A* and *B*. *C* does not get the share that *A* and *B* receive. Hence, it may in turn find it more difficult to secure co-operation with *D*. But there is no reason why *B* and *C* could not come to some sort of agreement. It may be true that *B* will not manage to elicit a high level of co-operative gain from *C*. But it is not implausible that the benefits from having a greater share of *X* still outweigh the reduced benefits from co-operation with *C*.

To put it differently, *A* and *B* will each calculate the advantages of having full access to *X*, compare these to the expected benefits of full inter-generational co-operation –

call these Y – and opt for the most attractive arrangement or bundle of X and Y they could possibly obtain. Since it is clearly impossible to fully deplete and maximise co-operative benefits at the same time, each generation will trade X off against Y until an equilibrium is reached.

Now, theoretically, it is still a possibility that this optimum converges with intergenerational equality. But Gauthier's approach could only guarantee this convergence by attaching an unrealistic weight to the relative benefit of intergenerational co-operation vis-a-vis the benefit of resource depletion. His argument suggests that present generation A should be prepared to give up its monopoly over a valuable resource, and satisfy itself with $1/n^{\text{th}}$ share of X , all in return for B 's co-operation. Now, it may be plausible, perhaps, that A would give up half or even two-thirds of its share in X to obtain a certain share of co-operative benefit in return. But is it also plausible that A would consider a drop to having a mere $1/10^{\text{th}}$ or $1/20^{\text{th}}$ or even $1/50^{\text{th}}$ share?³² It is not unreasonable, I think, to claim that generational overlap explains how just distributions between a limited number of generations could emerge. This claim seems quite credible for generations that directly overlap, like A and B , and possibly even for the next one or two generations. What is unreasonable is Gauthier's further claim that 'indirect co-operative links' extend 'to the farthest future of our kind'. For some n , the present generation will simply find it more profitable to hang on to what it now has and to break free from Gauthier's chain of inter-generational co-operation.

4.2 *The Morality of Feeling*

Now, Gauthier claims that the alternative routing to justice between generations invokes what he describes as 'the morality of feeling'. This morality, we may surmise,

³² For this and other objections, see L. S. Temkin, 'Justice and Equality: Some Questions About Scope', *Social Philosophy and Policy*, Volume 12, Number 2 (1995), pp. 72-104; K. Sauvé, 'Gauthier, Property Rights, and Future Generations', *Canadian Journal of Philosophy*, Volume 25, Number 2 (1995), pp. 163-176; A. De Shalit, *Why Posterity Matters: Environmental Policies and Future Generations* (London: Routledge, 1995), p. 96.

recognises that people normally take an interest in the well-being of their immediate successors. Other writers do indeed advance this solution. In *A Theory of Justice*, for example, Rawls claims that the parties in the original position will care about the well-being of at least some persons in the next generation.³³ We should regard them as 'heads of families' and 'having a desire to further the welfare of their nearest descendants'. Again the suggestion is that the 'overlapping interests' between two adjoining generations expand into a much longer chain of interests. *A* wants to benefit *B*, and *B* wants to benefit *C*, so *A* indirectly wants to promote the welfare of *C*. Now, this argument is I think still a form of individualism. It merely extends its motivational assumption by assuming that people are not completely disinterested in each other. *A* finds a reason to conserve because it has an indirect interest in the well-being of *C*. However, this also forms the weakness of the argument. The interests of *C* have weight, not as such, but only *because* they affect *A*'s own well-being. Moreover, it only acquires this weight through the intermediary position of *B*. It is obvious, therefore, that the interest of future generations could only have an impact in a diluted form. Kavka rightly remarks that this argument succumbs to a 'multiplier effect over generations'.³⁴

Ernest Partridge argues along similar lines, but in favour of a bolder revision of the motivation assumption.³⁵ Whereas Rawls limits himself to the view that we should think of ourselves as fathers or mothers, Partridge writes: 'healthy, well-functioning human beings have a basic and pervasive need to transcend themselves; that is, to identify themselves as a part of larger, ongoing, and enduring processes, projects, institutions, and ideals.'³⁶ People need to extend their interests and concerns beyond the narrow confines of their own banal lives, which are all too often ruled by 'gadgets', 'diversions' and 'opulence', and also beyond their own births and deaths. For only by

³³ J. Rawls, *A Theory Of Justice*, p. 128 and pp. 284-293. See also D. C. Hubin, 'Justice and Future Generations', *Philosophy and Public Affairs*, Vol. 6, Issue 1 (1976), pp. 70-83.

³⁴ G. S. Kavka, *Hobbesian Moral and Political Theory*, p. 444.

³⁵ E. Partridge, 'Why Care About the Future?', in: E. Partridge (ed.) *Responsibilities to Future Generations: Environmental Ethics*, pp. 203- 220.

³⁶ *Ibid.*, p. 217.

getting involved in such things as ‘philanthropy’, ‘the arts’, ‘social work’, and ‘political issues’ are we capable of leading truly meaningful lives. Only in these activities do we find an escape from ‘alienation’ and ‘narcissistic disorder’. Since Partridge also believes that the good of posterity is ‘a valid object of transpersonal devotion’, the convenient upshot is that ‘life is immediately enriched by the collective agreement of the living to provide for the well-being of the unborn.’³⁷ Hence, his theory is still meant to fit into the mutual advantage mould. Why care about the future? The answer is that in doing so I am in fact promoting *my own good properly conceived*.

However, there are important difficulties with this view. To begin with, I cannot help but feel deep uneasiness about the fact that these so-called self-transcendent concerns are somehow still grounded in the interests of the self. Partridge probably shares this worry. For he goes out of his way to show that ‘full’ self-transcendence requires that we engage with these concerns not because we thereby fill up our own existential voids or triumph over our mortality – this would quite obviously be self-defeating – but because we are under the (false) impression that these things are valuable for their own sake. So while transcendence is good for me, this cannot logically be my internal motive. Nevertheless, since Partridge looks at the issue from an external standpoint, for him the real (and necessarily secret) reason why I should care about the future is still that I have reason to care about myself.

Even so, I still find this attempt to bring everything together in self-interest highly questionable. First, there is the obvious suspicion that the whole approach is rigged from the outset.³⁸ From the initial assumption of mutual disinterest, our motivations are simply stretched up until they generate the correct results. The individualist first tries out the appeal to parenthood. Since that idea does not work, he introduces the notion of

³⁷ Ibid., p. 217.

³⁸ See Barry’s criticism of Rawls in ‘Justice Between Generations’, in: *Liberty and Justice: Essays in Political Theory 2* (Oxford: Clarendon Press, 1991), pp. 242-258, at p. 253. There he writes: ‘Our limited sympathies towards our successors are fed into the sausage-machine of ‘justice as fairness’ and returned to us duly certified as obligations.’ In the end, Rawls is left looking like ‘a conjurer putting a rabbit in a hat, taking it out again and expecting a round of applause.’

self-transcendence. And if that were insufficient too, I suppose, he could continue bringing in alternative assumptions until the desired outcome were reached. Now, the only way to remove the suspicion of rigging is to argue that self-transcendence is not just something that certain people do, but a constitutive part of the human good objectively speaking. If pressed hard enough, I might even come to accept this view. What I find much harder to accept, however, is that this good must necessarily involve a concern with the future. As Kavka notes, not all other-regarding interests have something to do with the good of posterity.³⁹ I might just want my football team to win the championship, study naval history, or advance the good only of my children. Why, then, assume that precisely this concern for the good of posterity is what enriches our lives?

Second, there is the further suspicion that Partridge's theory does not really fit the Hobbesian mould after all. Why care about the future? In the end, the answer is highly ambiguous. On the one hand, he claims that we should care about the future because posterity matters for its own sake. This describes what we actually have to believe and think in order to attain 'full' self-transcendence and thus our own personal good. However, oddly enough he does not appear to hold that this motive or attitude forms a bona fide reason for caring about the future. The real reason why we should care is that in doing so we are in fact helping ourselves. I find this position baffling. To all intents and purposes, we need to think like Kantians. We have to go native, as it were, and think of our fellow men as ends in themselves. But externally, that is to say, philosophically, this attitude only recommends itself as 'a more prudent policy for achieving self-enrichment and personal satisfaction'. What I am particularly unclear about is why Partridge thinks he still has to appeal to this prudential element. The standard Hobbesian account is clear enough in this respect. Prudence is supposed to give us a non-moral reason for being moral in the first place. So the crucial task is to

³⁹ See Kavka, *Hobbesian Moral and Political Theory*, p. 445.

show that moral conduct pays off in terms of our own happiness or good, *terms other than morality itself*. This seems a reasonable challenge. However, for Partridge this cannot be the case. After all, for him people are already supposed to think like fully committed Kantians. Hence, our own personal good cannot have an independent motivating or reason-giving force. In order to achieve 'full' transcendence, the philosophical truth about what we have reason to do must remain consigned to the crevices of the mind. But if this is his view, it makes more sense to embrace Kantianism in a more open fashion. The challenge is then to show that it pays off to care about the future *in terms of morality itself*.

5. What, if Anything, is Wrong With Resource Depletion?

The ethical individualist may therefore renew the Kantian objection. Leaving our motivations aside, moral individualism could still explain why conservation is the moral choice, or so this objection goes. The conclusion that the interests of future generations *ought* to count no less than our own is well within the reach of contractualist theories of justice. One key assumption of contractualist theory is that each individual qua individual enjoys the moral right to a reasonable justification for the institutions and policies that affect his or her well-being or the conditions thereof. This justification needs to be based on principles that no agent motivated to reach general agreement could reasonably reject.⁴⁰ When applied to generations, the contractualist assumption is that living generations need to be able to justify relevant policies and decisions to future generations. If Generation *A* wanted to deplete resource *X*, it could do so only on a principle that Generations *C* could not reasonably reject.

Let us say that *A* prefers depletion. Does *C* have a reasonable complaint? To put it differently, can *A* justify its preference for depletion to future generations? On first

⁴⁰ This is the gist of the contractualism of T. M. Scanlon, *What We Owe to Each Other* (Cambridge MA: Belknap Press, 1998).

inspection, this may seem highly unlikely. Depletion appears to rest on the principle of generational egoism at best. Surely, Generation *C* could reasonably reject this principle. Instead, it could offer the egalitarian principle that, all things equal, all generations have an equal claim to basic resources like *X*. This principle may further be grounded in the luck-egalitarian intuition that from the moral point of view it is ultimately arbitrary that we belong to generation *A* rather than *C*. No person has earned his place in the generational chain on the basis of ambition. As far as we know, all this is a matter of brute luck. Thus, no generation can derive a right to appropriate more than its equal share merely because it comes earlier than others. If this egalitarian principle goes through, then it seems easy to show that *A* violates the principles of justice. To be sure, it may be inefficient to give each generation an equal but infinitesimally small share of a finite resource. But the *right* to an equal share means that if present generations appropriate more than what is rightfully theirs, future generations at least ought to be compensated for their loss. They should be able to share in the efficiency gains. This seems to be the view that Brian Barry supports. According to him 'depletion should be compensated for in the sense that later generations should be left no worse off [...] than they would have been without the depletion.'⁴¹

But we do not have to be luck-egalitarians to make this point. We could also turn to Locke's theory of property. Developing this argument, Wolf suggests that while future persons have no property rights, their interest should weigh notwithstanding.⁴² For an explanation, he turns to the proviso, which stipulates that appropriation is legitimate only if it leaves 'enough and as good' for others. Wolf acknowledges that this formulation of the proviso is out of date. A key problem is that once we start taking future people into account it would be impossible to leave 'enough and as good'. To avoid this problem, he further suggests that we relax the proviso by linking it with

⁴¹ B. Barry 'Intergenerational Justice in Energy Policy', in D. Maclean and P. G. Brown (eds.), *Energy and The Future*, Totowa: Rowman and Littlefield, 1983), pp. 15-30, at p. 23.

⁴² C. Wolf, 'Contemporary Property Rights, Lockean Provisos, and the Interests of Future Generations', *Ethics*, 105 (July 1995), pp. 791-818.

Mill's principle that 'people should usually be left free to do as they wish provided that their actions are no[t?] harmful to others.'⁴³ He then deduces that 'A's appropriation of an unowned resource X constitutes a valid property claim iff no other person is harmed by A's appropriation of X.' In other words, we do not literally have to worry about leaving 'enough and as good'. We merely have to make sure that our initial appropriation, all things equal, does not impose a significant harm on others. Once this principle is up and running, we could offer yet a further explanation for our duties to posterity. Generation C has a reasonable complaint because A's appropriation of X constitutes a clear violation of the proviso.

However, it remains to be seen whether or not these principles really support the right results. The intergenerational context is commonly thought to cause at least four special complications. While the first two are not directly fatal, the third and fourth are more profound.

The first problem stems from the fact that we need to treat future persons as the bearers of interests and rights in the present. De George argues that something that does not exist cannot have anything. "This follows from the briefest analysis of the present tense form of the verb 'to have'."⁴⁴ Macklin argues that we cannot assign rights to future generations because the latter have 'no identifiable members'.⁴⁵ And pointing out that future persons do not presently have a capacity for choice, Hillel Steiner arrives at the same conclusion.⁴⁶ That said, there are I think persuasive arguments to the opposite. For example, Feinberg argues that we should concede that the interests, and therefore

⁴³ Ibid., p. 804.

⁴⁴ R. T. De George, 'The Environment, Rights, and Future Generations', in: K. E. Goodpaster and K. M. Sayre (eds.) *Ethics and Problems of the 21st Century* (Notre Dame: University of Notre Dame Press, 1979), pp. 93-105, at p. 95.

⁴⁵ R. Macklin, 'Can Future Generations Correctly Be Said to Have Rights', in: E. Partridge (ed.) *Responsibilities to Future Generations: Environmental Ethics*, pp. 151-155.

⁴⁶ H. Steiner, 'The Rights of Future Generations', in: D. Maclean and P. G. Brown (eds.), *Energy and The Future*, pp. 151-165.

the rights, of future persons are contingent on the existence of the latter.⁴⁷ Consequently, a future person does not presently have an interest or a right in being created. He also concedes that the identities of future persons are unclear. Nevertheless, we can still realistically assume that *someone* will exist in the future. It is sufficient to know that there *will be* an interest, even if we do not precisely know whose interest it will be. Nor should we be fazed by the fact that future persons are currently unable to exercise a capacity for choice. After all, this fact does not stop us from awarding rights to babies and the mentally ill. As Scanlon points out, people lacking the ability to form and express their will might not be moral agents in the full sense, but when they can be said to have an interest or a good, they could still be represented by trustees acting and deciding on their behalf.⁴⁸

The second complication arises from the fact that the preferences of future generations could quite conceivably be different from ours.⁴⁹ *Autres temps, autres mœurs*. It could be said that the future is like a foreign land or culture. Consequently, even if we wanted to, we simply could not know how to promote the good of posterity. Could we really be sure that future generations will appreciate Venetian architecture? Should we preserve what we think are precious landscapes rather than build more airports? Perhaps not, but I suspect that this argument does not carry enough weight to dismiss the above principles. In fact, it is unclear why the issue should be anymore problematic in the *intergenerational* context than in the *intragenerational* context. Contractualist approaches already rely on a thin theory of the good, a theory of primary goods or basic resources that facilitate any reasonable conception of the good. We may not know to the last detail what the interests of future generations are, but is it really

⁴⁷ J. Feinberg, *Rights, Justice, and the Bounds of Liberty*, (Princeton: Princeton University Press, 1980). See R. Elliot, 'The Rights of Future People', *Journal of Applied Philosophy*, Volume 6, No. 2 (1989), pp. 159-169.

⁴⁸ T. M. Scanlon, 'Contractualism and Utilitarianism', in: A. Sen and B. Williams (ed.), *Utilitarianism and Beyond*, pp. 103-128, in particular pp. 113-115.

⁴⁹ See A. De Shalit, *Why Posterity Matters*, pp. 102-103; T. Ball, 'The Incoherence of Intergenerational Justice', *Inquiry*, 28:3 (1985), pp. 321-337.

that presumptuous to predict that future generations will benefit, for example, from inheriting a world fit for human habitation, a stable institutional framework and a prospering economy? In other words, we may have to accept that there are limits to what we could know about the preferences of future people, but to suggest that we could know nothing seems implausible.

I should make clear that I am not assuming these objections to be either individually or even jointly decisive. That said, a more serious obstruction is caused by the third complication. As Laslett and Fishkin have remarked, future generations are potentially unlimited in number.⁵⁰ Compared to a thousand generations to come, our own well-being appears infinitesimally small. Consequently, the demands of egalitarianism are unrealistically high, unless, that is, we could somehow justify a discounting rate for future interests. But a discounting rate seems to be in breach of the core assumption of impartiality.

The fourth objection, which I want to discuss at greater length, is often described as the 'non-identity problem'.⁵¹ To see why, let us return to the example of the depletion of resource X. Thus far, we have assumed that the identities of the individuals belonging to generation C remain unaffected by A's choice. Parfit has argued that we should eliminate this assumption.⁵² In that case, the choice for either depletion or conservation will co-determine which persons, out of a larger domain of potential future people, will become the people that actually form Generation C. If A decides for depletion, C is formed by a particular set of people *dC*. However, if A decides for conservation, C consists entirely out of a set of different people *cC*. This difference is explained by the

⁵⁰ P. Laslett and J. S. Fishkin, 'Introduction', in: P. Laslett and J. S. Fishkin (eds.), *Justice Between Age Groups and Generations*, p. 6. See also T. T. Thompson, 'Are We Obligated to Future Others?', in: E. Partridge (ed.), *Responsibilities to Future Generations* (Buffalo/New York: Prometheus Books, 1981), pp. 195-202.

⁵¹ See D. Parfit, *Reasons and Persons*, Chapter 16; T. Schwartz, 'Obligations to Posterity', in: R. I. Sikora and B. Barry (eds.), *Obligations to Future Generations* (Philadelphia: Temple University Press, 1978), pp. 3-13; D. MacLean, 'A Moral Requirement of Energy Policies', in: D. Maclean and P. G. Brown (eds.), *Energy and The Future*, pp. 180-197; J. S. Fishkin, 'Limits of Intergenerational Justice', in: P. Laslett and J. S. Fishkin (eds.), *Justice Between Age Groups and Generations*, pp. 62-83.

⁵² D. Parfit, *Reasons and Persons*, pp. 351-355.

impact that this decision has on the routines and actions of the members of Generation A. Small differences may have big consequences on the long term.

Once this assumption goes through, are the complaints of future generations really that reasonable? As Parfit points out, Barry's egalitarian principle presupposes that without depletion *dC* would have been better off. But the truth of the matter is that without depletion they would not even have existed. If *A* opts for conservation, the future belongs to *cC*. Assuming that *dC* would not have been better off not existing at all, perhaps *dC* ought to be grateful for the selfishness of those presently alive. At any rate, it is no longer obvious that future generations have reason to protest because depletion is bad for *them*. Wolf's proviso is bogged down by the same problem. It is simply not true that full appropriation is harmful to *dC*. Wolf simply brushes the non-identity problem aside in a footnote. 'Fortunately,' he says, 'few accept Parfit's important arguments, though many find them perplexing.' Perplexing they are indeed, but if correct, the proviso does not explain what is wrong with depletion. And if Parfit is mistaken, I would like to find out why.

Note that we may grant Feinberg that, contingent on their existence, future generations have interests or rights now, and that we should not harm these interests. This, however, is simply not enough to avoid the problem. The difficulty is not due to a lack of rights, but results from two other factors:

(1) *The Person-affecting Constraint on Morality*: The moral principles here advanced assume that an act or policy could be wrong only if this act or policy is worse for some particular person (or generation) than another act or policy.⁵³

For obvious reasons, this feature is closely aligned with the main tenet of moral individualism itself. In judging acts and policies we take the interests of continuous persons as our guideline, not the interests of time-sliced persons or even non-humanist criteria. If some act or policy is to be objectionable, at least someone must be worse off.

⁵³ Here I follow D. Parfit, *Reasons and Persons*, p. 394. See also Kavka's 'obligation principle' in 'The Paradox of Future Individuals', *Philosophy and Public Affairs*, Volume 11 (1982), pp. 93-112, p. 95.

(2) *The Rejection of the Same Persons Assumption*: Future generations consist out of a set of potential beings whose actuality is, all else equal, determined precisely by the very policies and decisions that need moral appraisal. In other words, we have given up the assumption that *dC* and *cC* are the same people.

It is this combination of assumptions that leads to the *Catch 22* that Parfit describes. If we deplete, we know perfectly well that *dC* will exist and we also fully agree that it may have interests now. But opting for conservation will not make any prospective member of *dC* better off. We might still think that depletion is wrong, but we cannot appeal to the interests and rights of future people in order to explain why. Depletion is not wrong because *dC*'s rights or any rights are violated, and so we need to look for other solutions and principles.

5.1 *The Veil of Ignorance*

One obvious suggestion is that the problem would simply disappear if we applied the veil of ignorance in a more rigorously Kantian way than Rawls himself claims is advisable. Rawls has always favoured what he calls 'the present-time point of entry'. This assumes that the parties in the original position are aware of the fact that they are contemporaries. Hence, the principles of justice are also biased to the present.⁵⁴ However, if it were possible to include parties from all generations, would we not be certain to arrive at more impartial judgements? Given the risk that the parties belong to *dC*, conservation would seem to maximise the position of the worst off.⁵⁵

Rawls rejects this view of the original position on at least two grounds. The first argument for retaining his present-time point of entry is that this seems more consistent

⁵⁴ J. Rawls, *A Theory of Justice*, p. 139; 292. In order to nevertheless arrive at a just savings principle Rawls used to prefer to amend the motivation assumption. More recently, however, his view has changed. In *Political Liberalism* (p. 274) he holds that under the present-time of entry interpretation the parties will adopt the correct saving principle 'subject to the further condition that they must want all *previous* generations to have followed it.' For a critical analysis of this solution, see R. Paden, 'Reciprocity and Intergenerational Justice', *Public Affairs Quarterly*, Vol. 10, Number 3, July, 1996, pp. 246-266.

⁵⁵ For this view, see D. A. J. Richards, 'Contractarian Theory and Intergenerational Justice', in: D. Maclean and P. G. Brown (eds.), *Energy and The Future*, pp. 131-150.

with the general role he allots to the circumstances of justice.⁵⁶ Briefly, he believes that the scope of justice is circumscribed by empirical conditions of reciprocity and equality. However, because the purely impartialist approach disposes of these circumstances altogether, this argument poses no real threat to a more inclusive approach. This view rejects the circumstances of justice in the first place.

The second argument is more specifically aimed at the issue of justice between generations. According to Rawls, the inclusive approach fails because it ‘stretches fantasy too far’.⁵⁷ While he does not explain at length why this is the case, I suspect that here he is nevertheless correct. Once we have rejected the same persons assumption, it should be recognised that there are two ways in which a purely impartial approach could set up the original position. To begin with, we could admit all persons who will live or have lived at some point in history. This means that the parties in the original position know for a fact that they will *actually* live. Alternatively, the original position could contain all persons who could *potentially* live at some point in time. So, the parties know that they might or might not live, depending, among other things, on the principles or policies actually chosen. However, both interpretations are equally problematic. On the one hand, it has been claimed that the idea that we should include only actual persons is founded on illogical reasoning.⁵⁸ If Parfit is right, we should assume that the very principles of justice that are adopted in the original position co-determine who will actually exist. But then we would have to fix these principles before the procedure could even commence. For example, if we say that Generation *dC* is represented behind the veil of ignorance, then we have already committed ourselves to depletion. Once it is assumed that the parties in the original position know that they will exist at some point in the future there is little left to say. On the other hand, the idea that

⁵⁶ See B. Barry, ‘Circumstances of Justice and Future Generations’, in: R. I. Sikora and B. Barry (eds.), *Obligations to Future Generations*, pp. 204-248; and more generally his *Theories of Justice*.

⁵⁷ J. Rawls, *A Theory of Justice*, p. 139.

⁵⁸ B. Barry, ‘Justice Between Generations’, pp. 254-255; D. Parfit, *Reasons and Persons*, p. 392; A. De Shalit, *Why Posterity Matters*, pp. 110-111.

we should admit all potential persons may seem to do greater justice to the theory. At least it avoids fixing certain outcomes in advance. Nevertheless, this interpretation seems a non-starter for a different reason. The parties in the original position must now decide on their very own existence, and to put it mildly, it is not at all clear on what rational basis they are exactly supposed to do that. Assuming that all parties want to avoid non-existence, how do we maximise the position of the worst-off? To be frank, the whole exercise seems rather absurd. Rather than doing its job as an 'expository device', here the veil of ignorance only works to complicate the issue even further.

5.2 Kavka's Solution

A further and more interesting suggestion appeals to the Kantian imperative that we ought to treat persons *and their creation* as ends in themselves, never as means to some other end. This argument is pressed in Gregory Kavka's treatment of the issue.⁵⁹ He compares the non-identity problem to a range of what he thinks are comparable cases, most notably:

The case of the slave child: Assume that slavery is legal. A slaveholder offers \$50,000 to a couple to first produce a child and then sell it into slavery to him. The couple did not have any prior plans to create a child. But since they are keen to buy a yacht, the extra money would be extremely useful. Should they decide to go ahead with it?

Most of us think that this would be terribly wrong, of course. But how could we argue for this conclusion? As in the depletion case, the difficulty is that we cannot appeal to the interests of the child. If the couple reject the offer, they will not have that child. Assuming that being a slave is still better than not being at all, the child would be better off if the couple accepted the offer, as would the couple and the slaveholder. We cannot, then, appeal to principles that fall within the person-affecting constraint, which could

⁵⁹ G. S. Kavka, 'The Paradox of Future Individuals', *Philosophy and Public Affairs*, Volume 11 (1982), pp. 93-112.

reject policies or acts only on the condition that these policies or acts harm the interests of some person.

Essentially, Kavka suggests we tinker with and stretch up the person-affecting constraint until we have found an acceptable result. He aims to find principles that could show that it is possible to wrong someone without actually harming him.⁶⁰ Two ideas appeal to him in particular. First, the agreement is wrong because it creates a state of affairs that is *intrinsically* undesirable. Even though it harms no one, a world with the slave child is inherently bad, Kavka claims, because it contains a 'restricted life'. That is, it contains 'a life that is significantly deficient in one or more of the major respects that generally make human lives valuable and worth living.'⁶¹ Second, the agreement is wrong because the couple use the creation of a person as a means to an end, which in this case is a yacht. According to Kavka, it is wrong to treat a person as a commodity, even if in doing so we make him or her better off.

Now, these principles may or may not explain our intuitions in the case of the slave child. But the problem is that we cannot rush to the conclusion that these principles are similarly applicable to the depletion case. In fact, the Kantian imperative seems only of limited value here. Kavka pithily acknowledges that the depletion case is unlike the case of the slave child in the following way. While it is obvious that the couple treat the creation of the child as a means to an end, this is not so in the case of future generations. The present generation does not exploit the creation of a future generation as a means to its own well-being. Rather, the fact that future generation might contain restricted lives is merely a *side effect* of its wish to exploit a certain resource. Now, Kavka supposes that this requires only minor adjustment. Depletion is wrong, he says, because 'the *fact* of creation is used as a means of cancelling the prima facie moral prohibition on the acts in question that derives from their undesirable effects.'⁶² To put it differently, present

⁶⁰ Ibid., p. 97.

⁶¹ Ibid., p. 105.

⁶² Ibid., p. 111.

generations exploit their powers of creation to get around what should be a perfectly obvious moral obligation. However, this modification is not without its problems. The main difficulty is that the principle now *presupposes* depletion to be wrong. Consequently, it can no longer do the work that it does in the case of the slave child. The Kantian argument seems to rely on a prior moral principle to explain that there is an obligation to cancel in the first place.

And so it is Kavka's first principle that carries the full weight of the argument. The claim is that depletion would lead to a future state of affairs that contained restricted lives, that such a state is intrinsically bad, and that we are consequently obligated not to create it. Notice that according to this principle we are not under a *positive* duty to produce the best life or any life at all. It merely claims that we have a *negative* duty not to create restricted lives. So, it maintains that it is inherently worse to lead a restricted life than no life at all. But is this plausible? One important difficulty with the principle is its vagueness. Everything depends on how the notion of a restricted life is fleshed out. According to Kavka's definition, a life is restricted if it lacks at least one of the basic elements that make a life worth living. For example, the life of a slave is restricted because it lacks autonomy. Other examples he gives are the lives of the seriously handicapped, of people 'struck down in the prime of life by illness' and of many of those 'living in a very overcrowded world'. He further adds that restricted lives will still be worth living on the whole, sometimes even more so than many non-restricted lives. The former, he says, are only 'on average' less worthwhile than the latter.

It has been argued that this definition is inadequate because, in this form, the principle produces many counterintuitive results.⁶³ Take Ludwig, a deaf person. Is it bad that he ever lives, even if his life is on the whole more fulfilling and enjoyable than the average life? It is one thing to claim that it is a bad thing that Ludwig is deaf – although apparently even might be controversial. However, it is quite another thing to say that

⁶³ See D. Parfit, 'Future Generations: Further Problems', *Philosophy and Public Affairs*, Volume 11 (1982), pp. 113-172, at pp. 120-121.

Ludwig had better not live at all. This is unquestionably absurd. Parfit mentions another example. Many people believe that having children is something that makes life worthwhile. Is it *intrinsically* bad that there are people who cannot have children? Again, this seems highly unlikely.

To avoid these results, could we not simply define the principle more restrictively? Kavka might have suggested that we focus on people who lack not one, but several elements that make life valuable and worth living. Let us say, then, that Ludwig is not merely deaf, but also suffers from a more serious handicap. This leaves him in a chronic state of physical agony, depression, despair, and so on. If we pushed Ludwig's case to ever higher levels of deprivation, then it would surely follow, at one point, that his life were intrinsically undesirable? Now, that may well be the case. However, this strategy falls short for a different reason. First, notice that by increasing the level of restrictedness, we are also decreasing the scope of the principle. We may now have to conclude that cases like those of the slave child and depletion do not contain sufficiently restricted lives. Second, there is also a point at which the appeal to the intrinsic badness of restricted lives becomes entirely superfluous. If Ludwig's life is restricted to the extent that it is no longer worth living, then we might as well appeal to principles that fall within the person-affecting constraint. Such a state of affairs would not only be *intrinsically* bad, it would also be bad *for* Ludwig. To put it differently, Kavka's principle is in threat of being flattened between a lower and an upper limit of restrictedness. At the lower limit, it loses its intuitive plausibility. At the upper limit, the principle loses its scope and efficacy.

I have said that Kavka tinkers with the person-affection constraint. Let me briefly mention a more radical departure from this notion. Kavka argues it is undesirable to create restricted lives, less desirable than not creating anyone at all. Parfit takes the next logical step. He claims that it is undesirable to create a restricted life if we could have

created a less or non-restricted life instead.⁶⁴ According to him we should compare the possible lives that we could create and produce the life or the state of affairs that contains the greatest measure of utility. Depletion is wrong, not because it is harmful to anyone or produces restrictive lives, but because possible future generation *cC* would be better off than *dC*. Similarly, we ought not to produce Kavka's slave child because a free child, presumably, would generally be better off than an enslaved child. To put it briefly, of all possible future states of affairs, we should create the one that maximises total or average utility.

It falls beyond the scope of this chapter to discuss this maximisation criterion at length. Suffice it to note that it rejects person-affecting constraints altogether in favour of an entirely impersonal morality. According to the latter morality, an act or policy would be wrong if it promoted other than optimal future states of affairs. Whether or not an act or policy is harmful to a person is neither here nor there. Rather, what counts is the level of total or average utility, regardless of the persons to whom this utility is attached. While such an impersonal approach may indeed be successful in avoiding the non-identity problem, it encourages other unattractive conclusions. One persistent worry is that this criterion is too demanding. It may suggest a duty to reproduce when this would lead to a higher level of future well-being.⁶⁵ Another persistent worry is that utilitarianism supports 'absurd' and 'repugnant' conclusions in the domain of population ethics.⁶⁶ But even if we would manage to avoid these embarrassments, liberal individualists like Kavka could not subscribe to impersonal morality. To ethical

⁶⁴ Ibid., p. 120. See more generally his *Reasons and Persons*, Part 4.

⁶⁵ G. S. Kavka, 'The Paradox of Future Individuals', p. 99.

⁶⁶ See D. Parfit, *Reasons and Persons*, Chapters 17 and 18. Briefly, what he calls the 'absurd' conclusion argues that we have to add more people to the world as long as this maximises total utility. What he calls the 'repugnant' conclusion claims that we may have to reduce the population until we have maximised average utility. Parfit admits that he cannot avoid these conclusions. Notice also that utilitarianism does not have to subscribe to impersonal morality per se. For an approach that claims to be utilitarian and person-affecting at the same time, see J. Narveson, 'Future People and Us', in: R. I. Sikora and B. Barry (eds.), *Obligations to Future Generations*, pp. 38-60. Narveson is right to claim that his person-affecting utilitarianism avoids these awkward conclusions. But his approach copes no better with Parfit's non-identity problem than egalitarian and harm-based solutions. See also the discussion in A. de-Shalit, *Why Posterity Matters*, Chapter 3.

individualism personal interests and harm are never neither here nor there. Persons are what fundamentally matters in morality. Ethical individualism needs to find a solution to the problem of generations that is consistent with that central assumption.

Conclusion

Chapter II argues that personal identity should be seen as a normative ideal, not an anterior concept aiming at the justification of a particular normative theory. While personal identity is relational, these relations should be seen to facilitate responsibility, obligation and prudence, and so restore the ethical unity of agency. Chapter III and IV argue that associations and groups could have their own personal integrity. Chapter V, VI and VII investigate whether or not corporate personality is plausible as a normative conception.

To this purpose, Chapter V introduced three hypotheses. The first thesis claims that the so-called background goods of culture and language are irreducibly social. However, I concluded that these goods could be accommodated by liberal individualism. The second thesis advances the proposition that associativism explains why we have reason to act morally. The first problem with this view is that mutual advantage contractualism offers a rival explanation. The second problem is that impartial contractualism denies that this refutes moral individualism at an ideal-normative level. Finally, the third hypothesis is associated with Burkean conservatism. Briefly, the central claim is here that moral individualists cannot account for the practical connectedness that we are commonly thought to have with generations past and future.

This last hypothesis derives its strength from the fact that both contractualist objections lose their validity in this context. To validate this view, I focused on the individualist attempt to extend to scope of distributive justice to generations. More

precisely, I looked at the Hobbesian school of thought that justice should benefit the *de facto* interests of all parties involved – what Brian Barry calls ‘justice as mutual advantage’. I also looked at the Kantian notion that we should conceive of justice from the point of view of a hypothetical baseline of equality, a point of view from where all bias towards one’s own interests is barred – what Barry refers to as ‘justice as impartiality’. Banking on a variety of strategies, theorists of both schools claim to be able to show that present generations have certain duties of justice to future generations.

I argue that these claims are mistaken. This failure can be traced to one common assumption. While justice as mutual advantage allows us to be biased to our selves, and justice as impartiality demands that we treat people equally, both approaches maintain that individual well-being is what ultimately matters in a reason-giving sense. Both varieties of contractualism operate within the person-affecting constraint. They assume that an act or policy could be wrong only if this act or policy would be worse for some individual than another act or policy. Whether actions, policies and institutions are just or unjust pivots on how they affect the interests and rights of individuals.

The hypothesis of conservatism, then, may be the most promising. While it is one thing to claim that ethical individualists could explain why we have reason to attach significance to culture, or to be moral or sociable, it is something else to say that we also have reason to care about future generations. That said, in order to demonstrate that the idea of corporate personality fares better in this latter regard we still have to construct a positive alternative. I turn to this task in Chapter VI.

CHAPTER VI

WHY CARE FOR THE FUTURE?

Man owes what he is to the union with his fellow man. The possibility of forming associations, which not only increase the power of those alive at the time, but also – and most importantly, because the existence of the association outspans that of the individual personality – unite past generations with those to come, gave us the possibility of evolution, of history.

Otto von Gierke – *Das Deutsche Genossenschaftsrecht*¹

Introduction

In the previous chapter I have argued that individualism is a ‘one-generation philosophy’.² We saw that individualist theories struggle to explain why we should consider the interests of future generations. One important point to emerge from this discussion is that these theories fail at two levels. First, insofar as agents take only their own preferences as reason-giving, they do not provide the agent with a proper motivation. Second, even when assuming – unrealistically – that agents make no distinction between their own interests and those of others individuals, ethical individualism still falls short. The problem is not just that people are not motivated to *do* the right thing, but also that we cannot even explain what the right thing *is*. The normative case for the idea of corporate personality is closely linked with this double shortcoming.

So far the argument has been entirely negative. How should we proceed? I have already argued that the concept of a person should be included within the field of study of normative theory itself. I propose, then, that we try to make headway, not by constructing a novel principle of justice, but by changing our conception of the person.

¹ O. von Gierke, *Das Deutsche Genossenschaftsrecht*, in: *Community in Historical Perspective*, A. Black (ed.), p. 2.

² As J. Gray once put it in *Beyond the New Right* (London & New York: Routledge, 1993), p. 136.

That is, we first have to switch to an alternative unit of agency, and then look for action-guiding principles.

Leaving non-humanist accounts aside, there are two options. As we already saw, we could change the unit of agency to the temporally discrete stages of a person, and therefore reject the person-affecting constraint. This is what Parfit proposes. According to him, whether or not an act or policy is harmful to a future individual is of no importance to morality. What counts is the level of total or average utility contained in a future world. The notion of a temporally extended interest no longer places limits on consequentialism.

This chapter develops the second option, which in a sense remains within the person-affecting constraint. Rather than dropping this constraint altogether, this option seeks to make progress, not by curtailing, but by extending the concept of the person. It suggests that we should turn to the notion of corporate personality. I develop this idea as follows. Section 1 introduces the notion of transgenerational association by merging the concepts of collective and future-directed intention. Section 2 concentrates on *intrapersonal* distributions of well-being over time, and makes a distinction between an *aggregative* and a *narrative* conception of the goodness of a life as a whole. Section 3 applies these conceptions to temporal distributions between generations, and argues that in doing so we overcome the problems that obstruct ethical individualism.

1. Transgenerational Association

Earlier I argued that for a person to be identical over time we need a particular attitude towards the past and the future. That is, we need to acknowledge that our earlier and later person-stages are rationally and morally integrated. First, we must have a particular attitude to the future, a plan or future-directed intention. Second, we should also assume a certain attitude to the past. We have to take future-directed intentions as

reason-giving. What is more, we should be prepared to acknowledge past actions and decisions as our own, and try to maintain a level of internal consistency between them. Precisely this is what is meant by integrity.

I have also argued that corporate personality requires two similar attitudes. First, we must share an intention. A collection of individuals forms a corporate person when individual intentions are interlocking in such a way that they are aiming to achieve a purpose *together*. Second, we need to think in a mode in which this collective intention is considered to be normative. In other words, the collective intention should provide the group with a standard of conduct.

I now want to add a temporal dimension to the notion of corporate personality. Clearly, associations are not frozen in time. Not only do they connect individuals to each other, they also connect individuals over time. This could be explained as follows. To begin with, it should be pointed out that a collective intention could also be future-directed. Collective intentions are usually defined not just over the present actions of several people, but also over what they could achieve together over a period of time. Thus, the identity of a corporate person over time depends on a collection of people CP , whose intentions cohere rationally such that they could be said to be aiming to achieve something *together* at some future point in time. So, CP_t needs to be connected with its own 'person-stages' $CPS_{t1}, CPS_{t2} \dots CPS_{tm}$ in the same way that individual persons are. Associations have their own narrative, their own plot and structure, their own beginnings and ends. They emerge, moreover, where individuals think in a mode containing collective intentions referring to the life of an association as a temporal whole or at least an important chunk of it.

The key is that the continuity of individual and association differs in that the latter could span several generations. For the continuity of an association it is not necessary that its membership consist out of the same individuals. It is not the case, clearly, that an association could have continuity only if the same set of people at t share the same

intention at t_n . We know that every day many British citizens die and many new citizens are born. Yet in spite of this permanent flux we think that the state retains its continuity. What is more important, clearly, is that old and new members share the same intention. It is more important, in other words, that the association retains its integrity.

To see how the idea of corporate personality copes with the problem of generations, let us assume, then, that Generations *A*, *B* and *C* form part of the same trans-generational association.³ So, *A*, *B*, and *C* form a bundle of intentions with a degree of integrity. That is, their intentions are normatively guided by a collective and at the same time future-directed intention in the sense that I explained above. Now, if we give this associative relatedness the kind of importance that the moderate claim demands for personal identity, then we can perhaps begin to understand how generations are significantly connected.

To begin with, I shall focus matters on the problem discussed in the previous chapter. Why should present generations take the interests of future generations into account? Consider again Parfit's depletion case. Most people agree that the interests of Generation *B* and *C* should count, and that *A* therefore ought to conserve rather than deplete this resource. The problem, however, was that we could not find a theoretical justification for this intuition. Ethical individualists could appeal to theories of distributive justice. But when we applied these theories to the letter, it appeared that *A* does not have a strong duty of justice to hand down resource to its successors.

Why am I claiming that associativism provides us with a better theory? Associativism gives us a justification that could be compared to the reason why I should

³ De-Shalit, in his book *Why Posterity Matters*, has recently revived the idea of transgenerational moral and cultural community. While there are clear similarities between his approach and the one set out in this chapter, two important differences remain. First, and most importantly, I conceive of transgenerational association as a corporate *person*. That is, I do not necessarily want to invoke the communitarian notion of a moral or cultural community that is constitutive of its members' identities. Second, in order to capture the ethical importance of future generations, I draw more explicitly on a comparison with future selves and a conception of the goodness of a life conceived as a continuous whole. See also A. De-Shalit, *The Environment: Between Theory and Practice* (Oxford: Oxford University Press, 2000), Chapter 4. For another approach to the same idea, see J. O'Neill, *Ecology, Policy and Politics: Human Well-Being and the Natural World* (London: Routledge, 1993), Chapter 3.

care about my own private future well-being. According to the moderate claim, my future well-being has a bearing on my current decisions and resolutions because I am connected with my future selves PS_{t1} , PS_{t2} ... PS_m in the manner of a person. I should not be indifferent to a future that holds some catastrophe in store because it will affect *my* future good. I might conceivably have less reason to care if this catastrophe takes place in thirty years rather than in three months time, but I do have reason to care. To put it differently, it is through this relatedness that I extend my hopes and fears into the future. Now, if the future interests of a person could still acquire normative relevance in our present deliberations, we need to consider one key question more carefully. Given that the moderate claim holds, how should we distribute costs and benefits over our lives, not as a matter of justice, but rather as a matter of the goodness of a life taken as a whole? I shall look at some answers to this question in the case of individual persons first and then investigate what happens when we put them into operation in the case of transgenerational associations.

2. Temporal Distributions Within Lives

Notice again that temporal distributions within lives are not normally thought to be constrained by the Paretian or any other egalitarian principles of justice. It is not generally believed to be irrational, let alone wrong, to incur costs now in order to make greater gains in the future. To put it differently, compensation between the person-stages of one life seems unproblematic. In fact, it would seem that many of our plans and ends presuppose compensation. Typically, they require that we give up certain things presently in order to reap the rewards later on in life. Yet we do not believe that such sacrifices are irrational. And we definitely do not believe that making such sacrifices is wrong or unjust. This belief, moreover, is not invalidated by reductionism

about personal identity. For the moderate claim still allows us to frame the issue in terms of the goodness of *one* life rather than justice between *multiple* selves.

But if contractualism is implausible, how then do distributions of costs and benefits over the stages of a life affect its goodness? What kind of criteria do we appeal to? The standard reply, of course, is the utilitarian view that we should not distribute at all, or only insofar as certain distributions maximise the total utility contained in a whole life. The goodness of a life, therefore, depends purely on the sum-total of utility achieved at the end of a life.

This reply touches on two very distinct issues. First, there is the question whether we should be *partial* or *impartial* to the temporal parts of our life. The utilitarian view shares with egalitarianism that each person-stage counts for one, just as much as each person counts for one in the social calculus or distribution. There is, in short, no room for 'pure time-preferences'. Whether the same benefit or cost materialises now, in the near future, or at some later time in the future should have no impact on our decisions. All else equal, we should be impartial to *when* goods or costs transpire.⁴

The second question is whether temporal distributions have a value independent of maximisation. As they do on the social level, both the egalitarian and the utilitarian advocate impartiality. But while according to the former it matters how costs and benefits are divided over a life, the latter maintains that we should just maximise. To the goodness of a life as a whole, it matters not that each person-stage gets a certain share; what counts is how much it adds up to in the end.

However, I think a more persuasive reply would in fact challenge both components of the standard view. Consider the legitimacy of time preferences first. To be more precise, consider a time preference, all too familiar to most of us, for what comes *nearer* rather than *later* in the future, relative to where we are presently. Now, while reductionism about identity may not undermine the validity of prudential reason

⁴ See J. Rawls, *A Theory of Justice*, pp. 293-298; T. Nagel, *The Possibility of Altruism*, Chapter 8.

altogether, the moderate claim does appear to justify a discount rate for the future.⁵ This is not to say that the mere fact of temporal distance means that the costs and benefits of later person-stages weigh less. Temporal distance is as irrelevant as spatial distance. However, what justifies the discount rate is the weakening strength in the connectedness between the present and person-stages in the far future. If the relatedness with my later person-stages fizzles out as we move further into the future, then so does the relative weight of my future interests and preferences in my present reasoning. The strength of this reason is mediated by the strength of the relation. And in most cases this means that we may or even should rationally prioritise the near future over the further future.⁶

Second, and I think more importantly, David Velleman has argued that the goodness of a life taken as a whole cannot be entirely reduced to a simple addition of temporally discrete well-being.⁷ Over and above mere addition, some independent distributive criterion comes in. This criterion, however, is not fully captured by egalitarianism either. Rather, it suggests that the good life is partly shaped by its progression over time, by what he refers to as ‘the overall structure of events’ and ‘their narrative or dramatic relations’. In other words, he suggests that there is a distributive aspect of the good after all, and that this is captured by how we evaluate the story or the plot of a life. At any rate, in judging the narrative coherence of lives we do not just add up temporally

⁵ D. Parfit, *Reasons and Persons*, Chapter 14 and Appendix F. In Chapter 8 of that book, Parfit explores a different argument against temporal neutrality. Briefly, this holds that a strict rejection of time preferences implies that we should not distinguish, not just between the *near* and the *later* future, but also between the *past* and *future*. However, intuitively (and on any conception of personal identity) it matters very much whether desires or benefits and costs are in the past rather than in the future. We are clearly biased towards the future, which is in itself a departure from temporal neutrality. Here, I shall leave that issue for what it is and focus on the bias towards the near.

⁶ Michael Slote, in his ‘Goods and Lives’, in: *Goods and Virtues* (Oxford: Clarendon Press, 1983), pp. 9-37, argues for an opposite time preference in favour of *later* over the *earlier* stages of a life. But when he distinguishes between these stages, he does not mean later and nearer relative to where we are presently. He believes that our lives can normally be divided in a number of phases like childhood, adolescence, middle age, and so on. Judged from an absolute point of view, that is, regardless of the phase we find ourselves in presently, it is the case, as Slote says on page 23, ‘that what happens late in life is naturally and automatically invested with greater significance and weight in determining the goodness of lives.’

⁷ J. D. Velleman, ‘Well-being and Time’, in: *The Possibility of Practical Reason* (Oxford: Clarendon Press, 2000) pp. 170-199. See also S. R. Wolf, ‘Happiness and Meaning: Two Aspects of the Good Life’, *Social Philosophy and Policy*, Volume 14, Number 1 (1997), pp. 207-225.

discrete moments of well-being. As Velleman remarks, we do not just say, Monday was a good day, Tuesday a bad one, Wednesday an uneventful one, and so forth, until we have covered all the days of a life. Instead, we need to look at how lives evolve and progress towards their end.

Take one of his examples:

Consider two different lives that you might live. One life begins in the depths but takes an upward trend: a childhood of deprivation, a troubled youth, struggles and setbacks in early adulthood, followed finally by success and satisfaction in middle age and a peaceful retirement. Another life begins at the heights but slides downhill: a blissful childhood and youth, precocious triumphs and rewards in early adulthood, followed by a midlife strewn with disasters that lead to misery in old age. Surely, we can imagine two such lives as containing equal sums of momentary well-being. Your retirement is as blessed in one life as your childhood in the other; your nonage is as blighted in one life as your dotage is in the other.⁸

Now, *ex hypothesi* these two lives contain an equal amount of added momentary well-being. Even so, there is clearly a sense in which we tend to evaluate them differently. Hence, Velleman makes the wider point that in our judgements on the issue we are also looking at 'trends' rather than just 'sums'. But his example also suggests a more specific idea, namely that a life characterised by an upward trend is, all things equal, better than a life typified by a downward trend.

The wider point that trends matter still falls suitably short of the egalitarian principle of justice that trade-offs between earlier and later moments of a life are strictly inadmissible, as if each person-stage formed a separate agent. On Velleman's view, such trade-offs are only constrained by the narrative aspects of the goodness of a life taken as a whole. So, on the one hand, his view allows that distributive trends have independent value. Pure addition does not account for the goodness of a life alone. But on the other hand, it is not committed to the more specific notion of an equal distribution *per se*. That is, the notion that well-being has to be distributed equally or the

⁸ J. D. Velleman, 'Well-being and Time', p. 58.

thought that inequality between person-stages is allowed only insofar as this maximises the position of the worst off.

Some people are of course drawn to this egalitarian view for the following reason. They prefer a life with a relatively consistent and unbroken level of well-being over a life with extreme peaks of delight and many extreme depths of despair, even though both trends average out at precisely the same level. Their view is that it is better to forgo the greatest joy if we can thereby avoid the greatest sorrow and come down somewhere in the middle.⁹ This seems a respectable and judicious position. For example, Sen believes that it is normally ‘not the case that no importance is attached to *intrapersonal* distributions.’¹⁰ More precisely, he thinks there is an intuitive basis for saying that the *intrapersonal* equality between the utility of the life-stages contributes to the overall goodness of that life. But surely – and this is my point – it does not give us the entire picture. For a strict interpretation of egalitarianism would surely rule out too much. It would altogether disallow the possibility of people incurring costs in the present in order to reap a reward or benefit later.

For example, consider the following two career tracks.¹¹ Track A involves accepting scantily paid work for the first five years. The job earns you equally little social esteem. But little by little you will move up the wage-ladder. Eventually, you will be vastly

⁹ The Romantics disagreed. Goethe famously wrote:

Blissful
And tearful,
With thought-teeming brain;
Hoping And fearing
In passionate pain;
Now shouting in triumph,
Now sunk in despair;—
With love’s thrilling rapture
What joy can compare!

From: Goethe, J. W. von, *Egmont*, Act III, Scene II, translated by Anna Swanwick, Vol. XIX, Part 3, The Harvard Classics (New York: P.F. Collier & Son, 1909–14; Bartleby.com, 2001. www.bartleby.com/19/3/).

¹⁰ A. Sen, ‘Utilitarianism and Welfarism’, *The Journal of Philosophy*, Vol. 76, No. 9. (1979), pp. 463-489, at p. 470-471.

¹¹ Velleman’s example tests how we appreciate upward as opposed to downward trends. This example tests how we appraise upward trends in contrast to an egalitarian distribution.

better paid and widely respected. Track B starts at a higher level. It involves accepting an averagely paid and respected job. But there is no prospect of promotion. You will earn the same wage your entire life. Assume that both tracks add up to the precise same sum of temporally discrete well-being at the end of your life. Now, if the only criterion is maximisation, we should be indifferent. If we are egalitarian, we must endorse the second path. But my and I think Velleman's intuition is that the first option is more preferable. It is better to start at a low level and to progress towards a much higher level.

If, then, we have to judge a life as a whole, there may indeed be good reasons to reject a great variance or inequality in temporal well-being. We might have reason to prefer more evenly distributed to wildly fluctuating lives especially if we thereby avoid moments of extreme pain or immense depths of despair. But there are clearly also many good reasons to welcome a degree of temporary restraint for the purpose of some later accomplishment. Put differently, there is an intuitively strong case for the idea that an upward trend – the advancement of a life towards an end or telos – adds to the value of a life taken as a whole, and so justifies a departure from egalitarianism.

What is the rationale behind the intuitive importance of upward trends? As Velleman notes, it is not that the later stages of a life naturally count for more than earlier ones.¹² If this were the case, the importance of trends could still be reduced to the additive criterion. We would just have to affix a greater weight to later person-stages and upward trends would routinely produce greater aggregates than downward or egalitarian trends. But why should the later stages of a life count for more in the first place?

One answer to this question is that later events can retroactively influence the momentary well-being of earlier person-stages.¹³ Assume that you opt for Career Track

¹² This is the view of time preferences that Slote defends in 'Goods and Lives'. But while he believes that this preference is allowed, he does not challenge the criterion of maximisation. Velleman, on the other hand, does not make the point that the utility of some person-stages weighs more than others, either in a relative or absolute sense. His point is that trends or distributions matter for reasons irreducible to quantities of temporally discrete utility.

¹³ See the discussion in J. D. Velleman, 'Well-being and Time', pp. 67-68. Another answer refers to the importance of anticipation and rising expectations. See for example J. Rawls. *A Theory of Justice*, p. 421.

A. If successful in later life, then your initial hardship will have paid off. And according to this view, that makes the temporally discrete experiences of earlier person-stages less bad than they were back then. Again, if this were the case, trends would after all be reducible to sums. Upward trends would contribute *ex post facto* to the well-being of earlier person-stages. But is such retroactive influence really possible?

Velleman thinks we should reject all such attempts to explain the value of upward trends in terms of their effect on the total sum of momentary well-being. His point is that there is a perspective of the goodness of a life as a whole that is entirely different from and not reducible to the maximisation criterion. From this standpoint, the unity of a life is understood in terms of what Bratman calls a plan or future-directed intention and MacIntyre famously describes as a quest.¹⁴ But consequently, the goodness of a life as a whole should equally be judged in terms of the successful completion of plan or quest. As MacIntyre puts the point:

The unity of a human life is the unity of a narrative quest. Quests sometimes fail, are frustrated, abandoned or dissipated into distractions; and human lives may in all these ways also fail. But the only criterion for success or failure in a human life as a whole are the criteria of success or failure in a narrated or to-be-narrated quest.¹⁵

That is to say, we need to look at what the past, present and future stages of a life will eventually achieve together in terms of an end-product; and we also need to look at the adversity and barriers that were conquered in order to get there.¹⁶

Now, from this point of view, the reason why upward trends have a greater value than downward or egalitarian ones is that the later stages in a life tend to retrospectively

¹⁴ I have this discussed this conception of personal identity in Section 5 of Chapter 3 above.

¹⁵ *After Virtue: A Study in Moral Theory*, pp. 218-219. See also C. Taylor, *Sources of the Self*, Chapter 2.

¹⁶ Nietzsche may have held an extreme version of this view. In *Twilight of the Idols* (p. 542), he wrote: 'For what is freedom? That one has the will to assume responsibility for oneself. That one maintains the distance which separates us. That one becomes more indifferent to difficulties, hardships, privation, even to life itself. That one is prepared to sacrifice human beings for one's cause, not excluding oneself. Freedom means that the manly instincts which delight in war and victory dominate over other instincts, for example, over those of 'pleasure'. The human being who has become free – and how much more the spirit who has become free – spits on the contemptible type of well-being dreamed of by shopkeepers, Christians, cows, females, Englishmen, and other democrats. The free man is a warrior.'

change the *meanings* of the earlier stages. To be more precise, the former allow us to interpret the latter as a prelude to bigger and better things to come and to judge earlier stages in terms of their contribution to achieving precisely these things. For example, the earlier setbacks in Track A are, as Taylor points out, 'redeemed' because they turn out to be a link in a chain of events leading to the fulfilment of a greater purpose and later happiness.¹⁷ Once you are indeed successful and well-regarded in your profession, the temporary misery by which progress was wrought seems to detract only little from the overall good of your life. In other words, if we consider person-stages to be rationally linked by a 'narrated or to-be-narrated quest', then the evaluative emphasis shifts away from the sorrow and bliss of separate moments towards the overall end to which your working life evolves and progresses.

Consider what happens when Track A goes wrong. I sacrifice my momentary well-being for the next five years in preparation of a career in a particular profession. I am about to reap some of the benefits. But instead I am fired and forced to accept a job in an entirely separate field. Assume that on balance my new profession turns out to offer me the same financial and social reward. Now, even though my new career provides as much temporally discrete well-being, there is a sense in which I am still worse off than I would have been otherwise. The reason is that my five years of hardship now seem like wasted time. Not only does this period contain many bad moments, it is also no longer connected to the later fulfilment of a greater purpose or my later happiness. Put differently, these years lose their meaning as a prologue to a finale. While professional success would have redeemed earlier deprivation, my failure causes my misery to be also utterly pointless.

Notice that this is not meant to imply that there are now actually more bad moments in my life. What happens later on makes no difference to what is experienced at an earlier moment. The fact that then-and-then I had a bad experience cannot be undone.

¹⁷ C. Taylor, *Sources of the Self*, pp. 50-51.

Consequently, from the standpoint of impartial maximisation it really is the case that once a moment of sorrow has come to pass, it is fed into the calculus and will blemish my life forever. However, the thesis is here that there are two distinct standpoints from which overall well-being should be assessed. Instead of trying to force these intuitions into one comprehensive utilitarian metric we should recognise that the goodness of a life taken as a whole is irreducibly plural. A person, Velleman says, ‘has two distinct sets of interests, lying along two distinct dimensions – his synchronic interests, in being well off at particular moments, and his diachronic interests, in having good periods of time and, in particular, a good life.’¹⁸ What this means, then, is that there are also two criteria for evaluating what self-interested persons should rationally do, each with its own authority. On the one hand, the ‘synchronic’ criterion suggests, with the utilitarian, that we should look at the temporally discrete well-being of our person-stages. However, we should now add that, within this perspective, the moderate claim implies a departure from the assumption of impartiality. That is, we have a strong justification for allowing a time preference for the near as opposed to the further future relative to where we are presently. On the other hand, the ‘diachronic’ criterion argues that we should do what promotes the good when the unity of life is understood as a quest for excellence. It emphasises the importance of upward trends or the successful achievement of a certain plan or objective at the end of one’s life.

To sum up, if we are trying to find out how the goodness of a life taken as a whole is affected by the temporal distributions of well-being over person-stages, there appear to be at least three criteria at work. The first criterion denies that such distributions have other than instrumental value. It argues for outright maximisation of the sum-total of the utility of the person-stages that form one life. The second criterion expresses an egalitarian concern with distributions. It claims that we should prevent moments of extreme anguish and suffering. The third criterion suggests that temporal distributions

¹⁸ J. D. Velleman, ‘Well-being and Time’, p. 71.

matter from the standpoint that views the unity of a life in terms of a plan or quest. This standpoint advocates the importance of upward trends.

3. Temporal Distributions within Associations

The crux of the normative argument for associativism is that these criteria apply in similar fashion to temporal distributions within transgenerational associations like states. Corporate personality opens up a mode of reasoning at the social level that individualism keeps confined to *intrapersonal* issues. It suggests that we tackle the generational problem by appealing to the criteria that govern the goodness of a life rather than those of contractual justice. Put differently, it suggests that we view generations like we view the temporal stages of one person.

This essentially is the Burkean view, which stands directly opposed to what Thomas Jefferson believed. The latter viewed each generation as a separate sovereign nation. As he put it:

... between society and society, or generation and generation there is no municipal obligation, no umpire but the law of nature. We seem not to have perceived that, by the law of nature, one generation is to another as one independent nation to another.¹⁹

If we view generations like Jefferson does here, the almost inevitable upshot is that the issue could only be framed in terms of contractual justice and equality, even if this invites the difficulties raised in the previous chapter. Even more problematically, perhaps, the generational issue must be compared to the problem of distributive justice between nations.

I suspect that there is indeed a widespread tendency to take it as a matter of course that both issues are the expression of what is in the end one and the same problem: how

¹⁹ T. Jefferson, 'The Earth Belongs to the Living', Letter to James Madison (September 6, 1789), in: M. D. Peterson (ed.), *Writings* (New York: The Library of America, 1984), pp. 959-964, at p. 962.

to extend the scope of justice or equality to a wider scale. More generally, it is perhaps distinctive of normative political theory after Rawls to want to bring within the purview of justice and equality an ever-increasing range of issues, problems and domains. The realm where these principles have no competence – we could call it freedom of association – has been forced on a steady retreat. But, predictably, this retreat leads to difficulties in cases where justice and equality are not cut out for the work, and thus tend to generate counter intuitive or even disturbing results.²⁰ The idea that we should feel compelled to bring generational relations under the authority of justice is I think another example of this mistake. We should realise that we only feel so compelled because we have been tricked into believing Jefferson's conception of generational sovereignty. And as long as this peculiar idea continues to slumber at the back of our minds, we will be hard pressed to see how we could ever think otherwise. We will carry on struggling to overcome paradoxes, absurdities and complexities like the non-identity problem, without realising that mutual advantage and equality are not the only available approaches. Generational relations are not necessarily like the relations between sovereign nations. The former are not by definition placed under the 'law of nature' without 'municipal obligations'. In fact, insofar as Jefferson claims logical exclusivity, his view is plainly false. The fundamental difference between sovereign nation-states and generations is precisely that generations could conceivably belong to *one* nation-state. This way of thinking is not just intuitively plausible but, more importantly, also logically possible. The idea of corporate personality shows that it is no less possible to see generations as the temporal stages of one association or community than it is to view the temporal stages of a life as part of one person.

²⁰ Notice that this is precisely what Parfit does. To wit, he first brings the relations between person-stages under the competence of justice and equality by means of the reductionist argument. However, paradoxically, he then claims that by so extending their scope, the principles of justice lose their weight and ultimately even collapse in a form of impersonal utilitarianism. See *Reasons and Persons*, pp. 332-345. Intriguingly, the argument suggests that this attempt to forever increase equality's scope to wider domains is ultimately self-defeating.

I suggest, then, that we reject Jefferson's conception and replace it with the Burkean view of the state as a transgenerational person. Once we change to this associative perspective, we have to admit, of course, that we are not making any claims about a moral duty of justice, at least not as Kantians like to think of it. The reasons or obligations we are talking about do not have the form of categorical duty. At the same time, we should also recognise that we are not making any claims about what narrowly self-interested agents have reason to do, at least not as understood by Gauthier and other followers of Hobbes. While the reasons I propose are agent-relative, they are also framed in terms of the good of a wider group of people. There is, then, a sense in which the idea of corporate personality overcomes the classic dichotomy between other-regarding duty and self-regarding interest. Unlike Kantian duty, it provides us with ethical reasons that are rooted in the motivations, interests and preferences that are constitutive of the self. Unlike justice as mutual advantage, however, corporate personality expands the notion of the self such that it includes not only present and future interests and preferences, but also the interests and preferences of certain other people. However we describe these reasons, the main argument is here that they avoid some of the problems that beleaguer the idea of intergenerational justice.

3.1 Society is Indeed a Contract

To begin with, consider again the problem that confronts mutual advantage theories of intergenerational justice. These theories rely on the guiding intuition that it is not right to transfer resources or goods from one party to another, unless *both* parties stand to gain from the transfer. Going back to the depletion case, it therefore has to be claimed that conservation leads to gains for the present and future generations. Ultimately, this is not a plausible claim. While conservation leads to serious gains for *B* and *C*, it is difficult to see why generation *A* should agree to such a policy, even though the sacrifice it involves may be trivial in comparison. The key problem with these theories

is that the scope for mutual co-operation between generations is limited. Present generations can transfer goods and resources to future generations, but there is little room for transfers in the opposite direction.

Conversely, corporate personality offers a view of intergenerational co-operation that is strikingly different. Justice as mutual advantage, on the one hand, conceives of society as a co-operative scheme between rational individuals, each with a *separate* interest. Like Jefferson, it views generations in a similar way, that is, as each having interests entirely separate from the other. This is precisely where the problem originates. Corporate personality, on the other hand, compares generational relations to co-operation between person-stages. We already saw that the interests of person-stages do not have a separate standing – this view, indeed, would merely induce madness. Instead, our temporally discrete interests are linked and governed by the idea of the goodness of a life taken as a continuous whole. To put it differently, while there is a real sense in which person-stages have to co-operate, this occurs under the normative guidance of a kind of collective intention. It is co-operation in a mode that does not include the Paretian principle. It is true, of course, that there is no room for a transfer of goods and resources from the future to the past. But this should not stop me, as a rational agent, from making such transfers from the past to the future. In the *intrapersonal* mode, co-operative schemes operate on the basis of the principles of the goodness of a life taken as a whole. I make such transfers for reasons that have to do with the achievement of certain goals at a later period or even the end of my life, or the avoidance of extreme suffering, or even the maximisation of temporally discrete well-being.²¹

If there were to be intergenerational co-operation at all, then it would have to be in this mode of agency. That is, it must be co-operation under the normative guidance of a shared and collective intention, not the mutual advantage of sovereign generations, each

²¹ K. Sauvé, in his 'Gauthier, Property Rights, and Future Generations', p. 175, similarly suggests that 'contractors want to be participants in a society that extends and improves into the future, rather than one that does not.'

taken to have an entirely separate interest. This is what Burke must have meant when he famously wrote that society was indeed a contract, though

... not only between those who are living, but between those who are living, those who are dead, and those who are to be born. Each contract of each particular state is but a clause in the great primeval contract of eternal society, linking the lower with the higher natures, connecting the visible and invisible world, according to a fixed compact sanctioned by the inviolable oath which holds all physical and all moral natures, each in their appointed place.²²

The Burkean social contract is unlike any mutual advantage contract because it extends through time and across generations. It does not deny that we form mutual advantage partnerships for ‘occasional interest’ or for things that are merely ‘temporary and perishable’. Justice as mutual advantage may well apply to many of our dealings with contemporaries. Even so, each generation is at the same time joined in a partnership to obtain greater things, things that cannot be achieved by one generation alone. Society, says Burke, ‘is a partnership in all science; a partnership in all art; a partnership in every virtue and in all perfection.’²³ His quarrel with the French Revolution was essentially that this contract should never have been overturned. However, my point is here merely that the generational contract is not governed by the Paretian idea – it cannot be – but by the principles of the goodness of a life taken as a whole.²⁴

Notice that these principles do not deny that the transitory interests of each generation could still matter in a non-additive manner. As suggested, it might also be true that the temporally discrete well-being of each person-stage still matter in a non-

²² E. Burke, *Reflections on the Revolution in France*, p. 85.

²³ *Ibid.*, p. 85.

²⁴ It is perhaps not surprising that Burke’s liberal critics sometimes fail to understand what he means in the passage cited above. For example, Don Herzog, in his article ‘Puzzling Through Burke’, *Political Theory*, Volume 19, No. 3 (1991), pp. 336-363, at pp. 341-342, thinks this passage contains ‘excess conceptual baggage’ and is ‘mysterious’ and ‘perplexing’. According to Herzog, Burke’s view of the social contract provides little else than a set of theological and metaphysical ramblings, which have consequently been shown to be erroneous by ‘the astonishing success of modern mechanism’. However, what Herzog, like before him Jefferson and Thomas Paine, in his *Rights of Man*, P. S. Foner, (ed.) (Secaucus: The Citadel Press, 1974), typically fails to comprehend is that the ‘modern mechanism’ of the social contract is at least in one sense a failure. While it may liberate us from the past, it may equally liberate us from the future. Burke, in contrast, saw this weakness and attempted to depict a contract according to which past and future generations were included in the deal.

additive manner. The point is merely that these interests matter in a less than overriding way. To put it differently, it is not the case that each generation could veto any proposed distribution. Should we avoid imposing terrible hardships on some generations, even if this is compensated by the even greater joy of others? On the diachronic view, I can imagine there being such reasons. Yet I cannot imagine believing in any strict sense that a generation should never be allowed to be worse off. That principle is clearly counter-intuitive, as the depletion example shows.

Take another example. Dahrendorf feared that former communist states needed to cross a ‘valley of tears’ in order to reach more democratic, just and affluent pastures.²⁵ Assuming he was right, imagine further that the first generation after the capitalist revolution is less well off than it would have been under communist rule, the second generation breaks even, the third reaches out of the valley, and all next generations start on the route to the wonderful summits of capitalism.

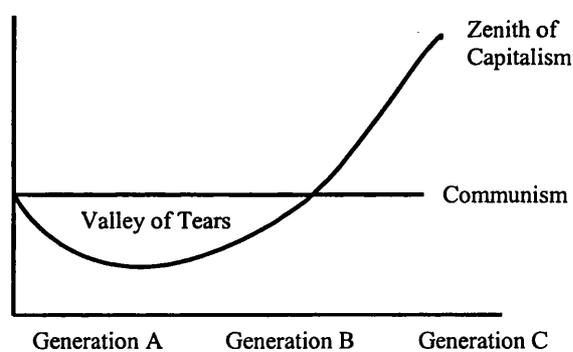


Figure 1 *The Valley of Tears*

We may ask: ‘Is this *right*?’ Mutual advantage theories must answer that it is not. The first generation receives nothing in return for its loss, even though that may be slight. It has been noted before that these theories are hostile to change – in this instance much more so than Burke’s – but they may also support communism. Should communism in this case really have prevailed? Clearly not, for we tend to ask ‘Is it worth it on the long

²⁵ R. Dahrendorf, *Reflections on the Revolution in Europe* (New York: Time Books 1990), p. 77.

run?’ And in asking this question, the unit of agency shifts from the individual towards the associative level. Here, from the synchronic standpoint we could claim that capitalism would add up to more well-being if we counted all generations. The diachronic standpoint compares trends and suggests that while communism perpetuates stagnation, capitalism will eventually elevate society to a higher stage of civilisation. We might claim, that is, that it matters more that these societies progress towards some measure of excellence. This may be an ideal of civil freedom or justice; it may be artistic grandeur; it may simply be a general level of well-being. But whatever its exact content, it is clear that in doing so we move away from co-operation in the mode of mutual advantage.

3.2 Why an Appeal to the Rights of Future People is Unnecessary

I turn to the problems of justice as impartiality. As we saw earlier, the sceptics of justice between generations often point out that it is hard to believe and ultimately illogical to claim that people who are not yet alive could have rights. Others deny this. If associativism is true, then we simply avoid this debate. The fact that generations are connected as described above means that the interests of future generations count as reason-giving for us now, just as our own future interests and preferences count as reason-giving on the moderate claim. We do not have to propose such rights because their interests are like our own interests. Associativism has the same impact as the moderate claim: while the latter explains why we have reason to care about my own future, the former extends the temporal scope of such concern beyond our own lifetime.

Perhaps more important is that we also steer clear of the non-identity problem. To recap, this problem stems from fact that future persons are merely potential beings whose actual existence in the future depends precisely on the actions and policies under scrutiny. In the depletion example, the trouble is that depletion is strictly not harmful to

anyone at all. Briefly, if we did not conserve, other future people would exist. Consequently, by switching from conservation to depletion we would not hurt anyone at all. However, if generations *A* and *C* were all part of one and the same association, then the problem would disappear. We know that *A* will be related to either *dC* or *cC*. Whichever generation will exist, it will form part of the same transgenerational association. *A*'s reason to conserve will thus derive from the good of the association.

It may help to make the following comparison. Notice that for the extremist, the non-identity problem occurs even within our own individual lives. For example, is smoking harmful to one of my future person-stages? I could argue that if I quit smoking, other future person-stages would exist. So quitting does not make any person-stage better off. Yet I do not think in this way precisely because the future person-stages that will actually exist will still be part of *my* life, whether I quit smoking or not. It does not matter that these are different person-stages. It is not important that I may or may not affect the interests of some temporally discrete person-stage. What matters is that both person-stages, if actualised, would become part of the greater whole that is my life. Thus, I have reason to quit if this has a beneficial effect on my life as a whole. That is to say, I can imagine having this reason because quitting maximises my total temporally discrete well-being. I can imagine having this reason because it avoids moments of extreme agony and despair, even though these moments are ultimately outweighed by the earlier enjoyment derived from smoking. What is more important, I can imagine having this reason because quitting allows me to successfully complete the pursuit of certain quests and ends later on in life. From this diachronic standpoint, what is particularly bad is that my life may be cut short in its prime. What matters are the unfinished projects and the things left undone.

The idea of transgenerational association allows us to view the issue of the temporal distribution over generations in the light of these three criteria. Generation *A* has a reason not to deplete because this would be hurtful to the association taken as a whole.

Thus, from the additive standpoint, *A* should not deplete if, all else equal, conserving turns out to maximise the sum-total of utility of all generations belonging to the transgenerational association.²⁶ From the diachronic viewpoint we might say that *A* should conserve if that avoids imposing enormous costs on one generation. But more importantly, we might also say that generation *A* should conserve because upward trends matter.

Consider another notorious problem. Assume that we could somehow phase out the human race in an entirely painless and voluntary manner. The present generation would unanimously agree not to reproduce. Humanity would vanish forever. But in return present people could deplete all resources free of guilt. Are there any reasons not to abort the species in this way? For causes similar to the non-identity problem, it is hard to think of a person-based answer. We do not harm anyone by deciding not to bring him or her into being. It is certainly not the case that individual rights are being trampled. However, there are such reasons when we turn to the idea of corporate personality. To understand why we should not abort the species, that is, we should change the unit of agency.

When Kavka asked himself this question, he admitted to believing that certain ‘collective enterprises of man’ – it seems he had the arts and sciences in mind – possess a value irreducible to their utility to us now. He found that ‘we should prefer a longer life of increasing accomplishment for mankind, to having human history cut short to facilitate present consumption.’²⁷ We should not abort humanity, he claimed, for the precise same reason that individual lives should not be cut short: we feel that their overall goodness depends on the progression and struggle towards the accomplishment of certain ends. A life cut short in its prime is often thought more tragic than death at a later stage, not so much because there will be less person-stages, and therefore less total

²⁶ If this were the only criterion, we would again encounter Parfit’s repugnant and absurd conclusions when considering population policy. See *Reasons and Persons*, Part IV.

²⁷ G. S. Kavka, ‘The Futurity Problem’, in: R. I. Sikora and B. Barry (eds.), *Obligations to Future Generations*, p. 197.

well-being, although that may play a role, but mainly because they tend to be lives spent in the pursuit of things never obtained, lives, therefore, with less narrative unity and too much time left unredeemed. If we look at generations as the stages of one association, it is for similar reasons that we should want humanity to continue.

Briefly notice that we now also avoid a number of further problems that are often associated with the notion of justice between generations. First, there is the difficulty that the preferences of future generations are unknown to us.²⁸ Consequently, justice as impartiality is limited to a thin theory of the good. De Shalit has pointed out that if history or the future is viewed, not like a foreign nation or country, but like a connected person-stage, then we must presuppose that future preferences and interests are to some extent continuous with ours.²⁹ So the argument does not have to remain entirely neutral to future conceptions of the good. We only have reason to care about the future precisely on the presumption that generations are in fact bound by a collective preference.

Second, there is the difficulty that future generations are almost unlimited in number and therefore dwarf the interests of present people. If we had to share resources with all persons equally and impartially, then the demands that the future places on us would be unrealistically high. In fact, given a finite quantity of resources, and an infinite number of generations, no one would get anything. However, like the moderate claim, associativism is less hostile to time preferences. Future interests only count as present reasons because of a kind of relatedness between person-stages or generations. Hence, we can now argue that future interests could legitimately be discounted insofar as the temporal relatedness diminishes. Put differently, according to the corporate model, the interests of distant generations could be assigned less relative weight than the interest of present or nearby generations, much in the same way that Hume proposed a social discounting mechanism.

²⁸ See Chapter V, Section 5.

²⁹ A. De-Shalit, *Why Posterity Matters*, pp. 102-103.

Third, there is the problem of motivation. I have argued that justice as impartiality does not explain why depletion is wrong. But even if this were false, justice as impartiality would still be vulnerable as it provides us only with agent-neutral reasons. In contrast to justice as mutual advantage, it does not show or pretend to show that conservation pays off in terms of *A*'s preferences. It argues instead that *A* ought not to deplete *in spite* of its preferences. The argument, typically, operates within an ideal-theoretical context where agents are imagined to fully comply with whatever impartial morality demands.³⁰ But it is unlikely to have a clear normative foothold outside this hypothesised context. Why be moral?

Associativism also fares better in this regard. It identifies reasons that are agent-relative, that is to say, grounded in the identities, preferences and interests of the agent, but argues further that such preferences and interests must be understood in a wider, collective sense. So it is able to explain why it pays off not to deplete by rejecting the categorical difference between self and other-regarding concerns. The benefit of this argument is that it is not restricted to the ideal-theoretical context that presupposes full compliance with the impartial standpoint. Assuming that we are connected to others in the way we have described, it will thus have greater normative purchase on our actual deliberations. To be precise, we have an argument that seems to appeal to considerations that are likely to actually motivate people: not pure impartiality, and not just the self, but also our relatedness to others.

It could be objected at this point that whether we have such reasons or not still depends on our being members of a particular transgenerational association. Do non-members have no reason, then, to concern themselves with future generations? To put the question differently, do rich states have no reason to concern themselves with the future interests of poor countries?³¹ We may not have to draw this conclusion. If there

³⁰ See J. Rawls, *A Theory of Justice*, pp. 8-9.

³¹ See for example Barry's critique of De-Shalit's communitarian approach in 'Sustainability and Intergenerational Justice', in: A. Dobson, *Fairness and Futurity* (Oxford: Oxford University Press, 1999),

were corporate persons, then we could fall back on theories of justice. Here, it is important to notice that the idea of corporate personality allows us to widen the person-affecting constraint on morality. As I argued in Chapter V, this constraint stipulates that an act or policy could be wrong if, and only if, that act or policy were worse for some individual than an alternative act or policy. If, apart from individuals, there were corporate persons, then an act or policy could be wrong if, and only if, it were worse for some particular *individual or association* than an alternative act or policy. So while depletion may not be worse than conservation for any individual, it may still be worse for an association, and therefore possibly unjust. Consequently, we have at least carved out some space to argue that non-members have moral reasons not to deplete resources if doing so would be harmful to a certain association. Similarly, from the perspective of international justice, we could also claim that transgenerational associations like states may have moral reasons to respect each other's future interests.³²

Notice also that I am merely arguing that there is room for such reasons. This room is created by the fact that the person-affecting constraint is now drawn wider than on individualist theories. Whether or not certain policies that are harmful to other associations are in fact unjust is a question that goes beyond the scope of the argument here. However, the answer could rely on familiar theories of justice and now apply them to the distributions between corporate persons rather than between individuals. Hence, we could argue, for example, that conservation offers co-operative gains to all associative parties involved. Alternatively, we could claim that each state or association has a basic right not to be harmed by others. In other words, the notion that collectives have continuous interests, and that these interests affect the wrongness of acts or policies, addresses some of the problems that perturb purely individualist approaches to

pp. 93-117, at p. 99. For a similar objection, see A. Dobson, *Justice and the Environment* (Oxford: Oxford University Press, 1998), pp. 106-107.

³² I have looked at the issue mainly from an internal perspective, that is, the standpoint of a member. But corporate persons may also have reason-giving weight from the external perspective of an outsider. As such, they may also impose moral claims on us. See here K. Graham, 'The Moral Significance of Collective Entities', pp. 21-42.

intergenerational distributions. While I find it difficult to see how moral individualism could avoid the conclusion that the notion of justice has no real intergenerational use, we might well return to this notion once it is recognised that there are corporate persons.³³

Conclusion

There is I think a sense in which Parfit's extremist *could* have been right. Normative reasons could have come exclusively in the shape of present aims or a preference or interest that we have currently. If so, our agency would have been temporally fractured. Every intentional episode would have been a separate agent. But there is also a sense in which this view *need* not be right. Our agency could still be temporally united by the normativity of future-directed intention. In other words, there is no argument that demonstrates the logical impossibility of temporally extended agency. Which view is true? This is a question of some importance, but, as I have argued, it is at bottom a normative question, and I have supposed that it must therefore be answered in reflective equilibrium with our basic intuitions, the practices that happen to be constitutive of our form of life, and the things that we tend to think are valuable to most humans. There may be a logically possible world in which Parfit's extremist is right, but, when we view the matter from this normative standpoint, it is not all that difficult to see that in this world, at least, he is not.

We have also seen that there is a sense in which ethical individualism *could* have been right. Our agency could have been fractured alongside our corporeal existence. If so, every individual human being would have been a separate agent, and this is often believed to be the case. But there is also a sense in which this view *need* not be right or at least not exclusively right. Our agency could still be united by the normativity of

³³ For the idea that we can avoid the non-identity problem by advancing collective rights see especially E. Page, 'Intergenerational Justice and Climate Change', *Political Studies*, XLVII (1999), pp. 53-66.

collective intention. As I argued in Chapter III, corporate personality is a logical possibility. Which view is true? Again, this is a normative question, a question that must be tested against our intuitions and the goings on that are constitutive of our world and lives. There may be a logically possible world in which moral individualism is exclusively right, but in Chapter V, VI and VII, I argue that this is not the case in this world.

A basic premise of my argument in these chapters is that we do not merely care about ourselves, or even about our contemporaries, but look back and forward in time, much further than our own lives. We are immersed not only in our own life-plans, but also in plans and projects with a greater history and continuity. In Chapter V, I argue that ethical individualism cannot account for this connectedness with the future. In this chapter, I argue that associativism offers a genuine alternative. Corporate personality allows us to compare distributions between generations to distributions between the stages of one person.

I shall next try to widen the scope of this argument. So far, I have based the entire idea of corporate personality on the intuition that future generations should have weight in our deliberations. This intuition may not be strong enough to carry this burden. Hence, I want to suggest a way in which this case may be further strengthened. Corporate personality, I propose, also explains why we are presently bound by rules, plans and commitments assumed by *past* generations. The idea that we could be bound by previous generations figures centrally in the debate on associative and political obligation. And it is to this debate that I turn in Chapter VII.

CHAPTER VII

WHY CARE FOR THE PAST

Wherefore, as any Man's Deed past is good as long as himself continueth: so the Act of a Publick Society of men done Five hundred Years sithence, standeth as theirs, who presently are of the same Societies, because Corporations are Immortal; we were then alive in our Predecessors and they in their Successors live still.

Richard Hooker – *Laws of the Ecclesiasticall Politie*¹

Introduction

Associations consist out of systems of roles, offices and positions, and whether anchored formally or not, these roles are partially, if not entirely, defined as duties, responsibilities and rights. Hence, associative members acquire all sorts of specific obligations to do this or that, depending on the specific sort of role or office they hold. For example, associates may be required to take special care of certain designated persons, to perform certain hazardous tasks, to be strictly impartial, and so forth. The specific and variegated content of such obligations is not what I want to discuss here. As I use the term, associative obligation refers more abstractly to the loyalty and commitment that all members owe to the association over and above the specific tasks attached to a particular role or office. In this chapter, I focus specifically on the question how political obligation could be justified or normatively grounded. In the context of this question, I want to further strengthen the case for corporate personality or associativism. As Heath Wellman has recently defined the term, associativism maintains that we should accord a basic and independent and obligation-generating weight to relational facts.² It claims that the mere fact that individuals are

¹ Cited in P. Laslett and J. S. Fishkin, 'Introduction', in: P. Laslett and J. S. Fishkin (eds.) *Justice Between Age Groups and Generations*, pp. 1-32, p.15.

² C. Heath Wellman, 'Relational Facts in Political Theory: Is There Magic in the Pronoun 'My'?', *Ethics*, 110 (2000) pp. 537-562.

interconnected in a specific way, that is, the mere fact of their membership, provides us with the justificatory grounds for their associative obligations. This associativist thesis or principle is often thought to have a specific impact on the following set of interrelated issues.

First and foremost, there is the long-standing problem of political obligation. Put simply, associativism claims that it is the fact of political membership or of citizenship that explains why we are obliged to comply with the law or have a wider obligation to protect and advance the interests of the political community.³ Second, there is the claim that the members of a political community have a *special* obligation to consider the needs and interests of their fellow-associates, roughly analogous to the special obligations that we have to our relatives or friends or neighbours. Here, associativism argues that such special obligations emerge out of the factual presence of such relations, without needing to invoke consent theory or a *general* duty applicable to humanity at large.⁴ Third, though less obviously, there is the idea that the members of a political association are in a collective sense morally accountable for the actions of their state. In this case, associativism argues that citizens partake in a collective sense of responsibility or guilt or desert merely on the basis of their *de facto* membership.⁵

With a few notable exceptions, most liberal theorists take exception to these claims for a number of reasons to be discussed later.⁶ But it is important to also note that liberalism, in contrast to philosophical anarchism, still supports the idea that political obligations are normally legitimate. Likewise, many mainstream liberals are also keen to argue that people have legitimate grounds to prioritise their fellow-citizens over the interests of humanity at large. The difference between the associativist thesis and these

³ B. Parekh, 'A Misconceived Discourse on Political Obligation', *Political Studies*, XLI (1993), pp. 236-251, draws special attention to this distinction.

⁴ For the terminology of 'special' and 'general' obligations, see H. L. A. Hart, 'Are There Any Natural Rights?', in: A. Quinton (ed.), *Political Philosophy* (London: Oxford University Press, 1967), pp. 53-66.

⁵ See M. Gilbert, 'Collective Remorse', pp. 123-140; 'The Idea of Collective Guilt', pp. 141-153, reprinted in her *Sociality and Responsibility*.

⁶ A notable exception is R. Dworkin, *Law's Empire*, Chapter 6.

broadly liberal approaches consists largely in the fact that for most liberals the mere occurrence of relatedness or the mere fact of membership does little in itself to establish the normativity of these special obligations. For most liberals, the legitimacy of special obligation depends entirely on the logical derivation of these obligations from the higher root of our universal and general duty to mankind.

Along these lines we can distinguish between two different justificatory theories, each aiming at the same objective. Associativism grants a special normative quality to relational facts. It suggests that there is something special about political relationships and bonds such that we can understand, among other things, why citizens have special welfare obligations to each other or why they owe a degree of loyalty to the state or even why they sometimes have to own up for what their state does. As Horton writes:

My claim is that a polity is, like the family, a relationship into which we are mostly born; and that the obligations which are constitutive of the relationship do not stand in need of moral justification in terms of a set of basic moral principles or some comprehensive moral theory.⁷

And in his book *On Nationality*, Miller argues for a view that ‘holds that relations between persons are part of the basic subject-matter of ethics, so that fundamental principles may be attached directly to these relations.’⁸

Most individualist liberals, however, tend to prefer what Wellman calls the ‘reductionist’ model. Here, I want to borrow Scheffler’s term ‘assimilationism’.⁹ According to this approach, relational facts have no basic obligation-creating significance. There is no magic in the pronoun ‘my’, as the view is often put. Special obligations are, like all moral concerns, ultimately grounded in a universal master principle, but that principle could be kneaded and massaged until it is capable of explaining why political obligation emerges and why we ought to have greater concern for the needs of associates than non-associates.

⁷ J. Horton, *Political Obligation* (Basingstoke: MacMillan, 1992), p. 150.

⁸ D. Miller, *On Nationality*, p. 50.

⁹ S. Scheffler, ‘Families, Nations, and Strangers’, in: *Boundaries and Allegiances: Problems of Justice and Responsibility in Liberal Thought* (Oxford: Oxford University Press, 2000), pp. 48-65.

I proceed as follows. Section 1 argues that at the heart of the problem of political obligation lies the question whether one generation could bind another. Section 2 introduces the assimilationist model and discusses both a general and a more specific objection to this model. Section 3 advances a value-dependent associativist alternative. Section 4 explains how this alternative tackles the most important objections to associativism.

1. Whether One Generation Has a Right to Bind Another

I want to start by following Hume in claiming that most of us are *born* into political obligation.¹⁰ As Hume pointed out, it may or may not be true that at some mythical moment in the past our ancestors expressly consented to a certain original contract, but it is an indisputable fact that we, who are alive presently, have never done so. If there are any political obligations at all, express consent has very little to do with it. Instead, political obligation comes to us as a bequest from the past. One way or another – how precisely is not at issue here – our ancestors developed structures of political authority and law, and in doing so they managed not only to bind themselves, but also every generation since. Fundamentally, the challenge is to show how such obligations can be passed on from one generation to the next. The central question of political obligation is, as Thomas Jefferson once put it, ‘whether one generation of men has a right to bind another’.¹¹

The claim that past generations obligate present generations is perhaps most visible in the following two cases. To begin with, there is the case of constitutional precommitment.¹² Constitutions are generally intended to put limitations on the future

¹⁰ David Hume, ‘Of the Original Contract’.

¹¹ T. Jefferson, ‘The Earth Belongs to the Living’, p. 959.

¹² S. Holmes, ‘Precommitment and the Paradox of Democracy’, in: J. Elster and R. Slagstad (eds.) *Constitutionalism and Democracy* (Cambridge: Cambridge University Press, 1988), pp. 195-240, and more generally S. Homes, *Passions and Constraint: On the Theory of Liberal Democracy* (Chicago: University of Chicago Press, 1995).

decisions of states. They form an attempt to obligate ourselves to follow one particular course of action rather than another. There are at least two reasons why constitutional commitment is thought to be controversial. First, there is the question whether a sovereign people could bind *itself* to some future undertaking. Traditionally, this has been a moot issue. It has often been thought that a sovereign power capable of binding itself at one time retains the power to release itself at later convenience.¹³ But what is more important here is that constitutions do not just bind the living; they also obligate or claim to obligate future generations. For example, the Preamble of the United States Constitution (1789) reads:

We the People of the United States, in Order to form a more perfect Union, establish Justice, insure domestic Tranquility, provide for the common defence, promote the general Welfare, and secure the Blessings of Liberty *to ourselves and our Posterity*, do ordain and establish this Constitution for the United States of America.¹⁴

Second, there is also the case of commitments to third parties. In one way, this case is less controversial. Once an individual or a state binds itself to a future undertaking involving the participation of another, it cannot quite so easily release itself from its earlier commitment. However, the intergenerational aspect of the problem remains. When a sovereign state acquires an obligation to another party this continues to bind all successive generations. Thus, Article X of the Treaty of Utrecht (1713) affirms:

¹³ In his sequel book *Ulysses Unbound: Studies in Rationality, Precommitment, and Constraints* (Cambridge: Cambridge University Press, 2000), Elster turns out to be sceptical of the idea of constitutional binding. Self-binding, he argues, depends on the presence of 'external institutions or forces' (p. 95). Ulysses could bind himself to the mast, making it *physically* impossible to give in to temptation, but '*there is nothing external to society*'. Thus, one reason why he objects to constitutional self-binding is its impossibility. However, Elster fails to appreciate that self-binding has *internal* aspects. That is, he fails to see that the planning and time-slice models of agency are *normative*. They prescribe what *rational* agents do. The planning theory, on the one hand, claims that rationality entails that we take earlier intentions as reason-giving. The time-slice model, on the other, argues that we ought to consider these intentions as external factors. To put it differently, both models advocate a different idea about what a rational attitude to the past consists in. To suggest that, in the absence of external ties, self-binding is impossible is already to assume the time-slice view. See E. F. McClennen, *Rationality and Dynamic Choice*, pp. 14, 231-238; S. J. Shapiro, 'Ulysses Rebound', *Economics and Philosophy*, 18 (2002), pp. 157-182.

¹⁴ *The Constitution of the United States of America*, 1787, in: C. Rossiter (ed.), *The Federalist Papers* (New York: Mentor Book, 1999), pp. 509-530, at p. 509. Emphasis added

The Catholic King does hereby, for himself, *his heirs and successors*, yield to the Crown of Great Britain the full and entire propriety of the town and castle of Gibraltar, together with the port, fortifications, and forts thereunto belonging; and he gives up the said propriety to be held and enjoyed absolutely with all manner of right *for ever*, without any exception or impediment whatsoever.¹⁵

As the example makes clear, the obligations and commitments of one state to another are not nullified as soon as one king makes place for another or when one generation is succeeded by the next. They are inheritable.

Do we have such obligations? Now, that one individual (or one generation) could and should limit *its own* future discretion is plausible if imposing such limitations would generate certain pay offs on the long run or would be normatively desirable for other reasons. At the level of the individual agent, it is this suggestion that informs Bratman's planning model.¹⁶ While intentions are in principle not irrevocable, Bratman reminds us that the cognitive capacities of humans are limited. Given these limitations, it is not prudent to have to reconsider each and every alternative before we could proceed to act. It is more rewarding to pre-commit to a basic plan as the default, and so to funnel deliberation in more precise directions. What is more, we need to commit ourselves to future actions in order to be able to attain ends or conceptions of the good that require co-ordination and co-operation between our person-stages. In other words, precommitment allows us to secure the long-term benefits.

At the collective level, Holmes makes a similar point about constitutional pre-commitment. Paradoxically, constitutional binding is 'enabling' rather than purely 'disabling'. For example, if we pre-commit to a set of decision rules and side-constraints, future deliberation could focus purely on solving problems of substance rather than having to engage in a costly re-evaluation of the procedures. Moreover, according to Holmes, it is only after such constitutional procedures and constraints have been put in place that we have the proper conditions for political deliberation to begin

¹⁵ Cited in: A. Heath, 'Research Roundup: The Assault on Gibraltar', *The European Journal*, Vol. 9, Number 3 (2001), pp. 8-9, p. 9.

¹⁶ See my discussion of his planning model of agency in Chapter II above.

with. That is, we need antecedent rules to conduct and organise the political process before political deliberation can actually commence. ‘The people cannot act as an amorphous blob.’¹⁷

However, this argument does still not show that constitutions could also obligate *future* generations. Jefferson was aware that liberal constitutionalism relied on the idea of corporate personality. He argued that it was wrong to allow the dead to obligate the living, either by constitutional means or by contracting debts to third parties. He grounded his case in his distinctive view of generations as sovereign agents. As he put it: ‘We may consider each generation as a distinct nation, with a right, by the will of its majority, to bind themselves, but none to bind the succeeding generation, more than the inhabitants of another country.’¹⁸ By viewing generations in this way, he made it perfectly clear why he thought placing constraints on future generations was illegitimate. What he described as a ‘perpetual constitution’ amounted to little more than perpetual oppression. Essentially, the earth belonged to the living, and each living generation enjoyed the right to a fresh beginning, unencumbered by the commitments and debts of the dead. As he put it, the dead had no rights: ‘They are nothing; and nothing cannot own something.’¹⁹ Hence, he insisted that constitutions, positive law and national debts naturally expired after nineteen years. This was how long it took for one majority to be replaced by a new generation, or so Jefferson reckoned.

¹⁷ S. Holmes, ‘Precommitment and the Paradox of Democracy’, p. 230. Holmes’s notion of enabling rules is similar to the Searle’s concept of constitutive rules set out in his *The Construction of Social Reality*.

¹⁸ T. Jefferson, ‘Debt, Taxes, Banks, and Paper’, Letter to John Wayles Eppes (June 24, 1813) in: M. D. Peterson (ed.), *Writings* (New York: The Library of America, 1984), pp. 1280-1286, at pp. 1280-1281. Paine, in his *Rights of Man*, arrived at a similar conclusion. He wrote (p. 56): ‘Those who have quitted the world, and those who are not yet arrived in it, are as remote from each other, as the utmost stretch of mortal imagination can conceive. What possible obligation, then, can exist between them; what rule or principle can be laid down, that two nonentities, the one out of existence, and the other not in, and who never can meet in this world, that the one should control the other to the end of time?’

¹⁹ T. Jefferson, ‘Reform of the Virginia Constitution’, Letter to Samuel Kercheval (July 12, 1816), in: M. D. Peterson (ed.), *Writings* (New York: The Library of America, 1984), pp. 1395-1403, at p. 1402. For the interests and rights of past individuals in liberal political theory, see T. Mulgan, ‘The Place of the Dead in Liberal Political Philosophy’, *The Journal of Political Philosophy*, Vol. 7 (1999), pp. 52-70; Brecher, B., ‘Our Obligation to the Dead’, *Journal of Applied Philosophy*, Volume 19, No. 2 (2002), pp. 109-119.

The argument that self-binding enables rather than disables does not as such address this crucial point. Holmes is probably correct to remark that such the introduction of such expiry dates would result in a major catastrophe. The anticipation of a constitutional repeal alone would probably suffice to propel society in a spiral of madness. To be sure, this is an excellent argument for *self*-binding. But the nub is that it is a bad argument for encroaching on the liberty of *others*.²⁰ And the nub of Jefferson's objection was that future generations were others. To put it differently, once we accept Jefferson's view, and still want to argue that our commitments bind future generations, then we need to look for other justificatory strategies.

The foremost candidate is of course the principle of consent. However, since no state relies on the express consent of its citizens, the only plausible candidate is the principle of *tacit* consent. If generations were sovereign, so the idea goes, then we should assume that each generation consents to the laws of previous generations by choosing to accept the property, goods and benefits it inherits from them.²¹ For example, if a new generation elects to remain in its country of birth, it thereby tacitly consents to its constitution and law, and it also inherits all earlier incurred debts. The underlying idea, of course, is that political obligations are attached to the property and the land that children expect to receive from their parents.

Now, this is not the place to look at consent theory at length. However, perhaps it will suffice to note that the theory of tacit consent relies on the condition that consenting generations should at least have the opportunity *not* to consent without facing

²⁰ Elster would seem to agree. He writes: 'It is only by the dubious device of construing the nation over time as a single entity that constitutions can be viewed as acts of *self*-binding.' In his 'Interpersonal Choice and Political Thought', in: G. Loewenstein and J. Elster (eds.) *Choice over Time* (New York: Russell Sage Foundation, 1992), pp. 35-53, at p. 37. In his more recent book *Ulysses Unbound*, p. 92, he further turns away from the initial idea that constitutions are correctly construed as mechanisms for *self*-binding. Apart from the objection that constitutions may not bind at all (see above), the problem is that 'constitutions may bind others rather than being acts of self-binding'.

²¹ As Locke proposed in his *Second Treatises of Government*, edited by C. B. Macpherson (Indianapolis: Hackett Publishing Company, 1980), especially Chapter VIII, paragraph 116, p. 62.

unreasonably high costs. And in the case of states the problem is invariably that this condition is not met. How could we possibly avoid inheriting the land and other goods from our ancestors? After all, we cannot simply choose to settle down elsewhere.²²

How, then, do generations put each other under political obligations? This is the puzzle that needs solving. To make headway, we could turn either to associativist or assimilationist justifications. I want to look at the latter first.

2. Impartiality and Special Allegiance

Essentially, the assimilationist model explains *political* obligation in terms of *moral* duty. It ultimately relies on the basic premise that the laws of morality are general such that they apply to all human beings equally and impartially. As Hart noted in his seminal article on the topic, general rights and duties are rights and duties that people have 'if they are capable of choice: they have it *qua* men, and not only if they are members of some society or stand in some special relation to each other.'²³ In other words, morality is impartial to all particular identities.

Two standard objections have been raised to this model. The first objection invokes a fairly wide-ranging scepticism pertaining to the basic premise of universality and impartiality. This objection argues that in order to get an operative moral theory up and running, we need to pack it with value-contents, which are invariably entrenched in our local *Sitten*. Important though this issue may be, this is not the place to deal with it.

Instead, I want to investigate the second objection, which Ronald Dworkin describes as the 'special allegiance objection'.²⁴ This objection points out that political obligations are essentially *special*. In contrast to moral duties, their scope is restricted. We have

²² Hume already pointed this out in his essay 'Of the Original Contract'. We simply cannot avoid inheriting their benefits, whether we want to or not. In this sense, past generations are like Nozick's public address system. See his *Anarchy, State, and Utopia*, pp. 93-95. See also R. Dworkin, *Law's Empire*, pp. 192-195.

²³ H. L. A. Hart, 'Are There Any Natural Rights', p. 53.

²⁴ Dworkin, *Law's Empire*, p. 193.

such obligations, not in virtue of our humanity, but in virtue of our membership of a political association. So Frenchmen only have political obligations to France, New Zealanders only to New Zealand and their fellow-countrymen, and so on. In this sense political obligations are comparable to family obligations. In the latter case, surely, the impartialist does not want to suggest that it is immoral for parents to nourish their own children rather than relieve suffering in general? Certainly, we should probably donate our old shoes, rags, and what not more to Oxfam, but the idea that we are morally compelled to treat everyone impartially is clearly absurd. As Williams puts the complaint: 'For morality the ethical constituency is always the same: the universal constituency.'²⁵

However, the by now standard response is that this objection rests on a flawed reading of what impartiality entails. The constituency of morality is not always universal. In fact, this response suggests that we should distinguish between two levels of impartiality.²⁶ First, 'substantive' impartiality demands that in our *actions* we betray no prejudice to people's particular identities and our relatedness to them. Evidently, there can be no compromise between impartiality and special allegiances here.²⁷ Second, 'formal' impartiality merely demands that the *principles* that are supposed to ground or justify our actual treatment of other people should be acceptable to all reasonable persons.²⁸ This is usually taken to mean, among other things, that these principles must at least be generic or impartial *in form*. For example, 'British Interests Come First' is not a principle that is likely to survive the test. However, at this justificatory level, impartiality does not mean that principles must be generic *in content*.

²⁵ B. Williams, *Ethics and the Limits of Philosophy*, p. 14.

²⁶ See for example B. Barry, *Justice as Impartiality* (Oxford: Clarendon Press, 1995); T. Nagel, *Equality and Partiality* (New York: Oxford University Press, 1991); A. Gewirth, 'Ethical Universalism and Particularism', *The Journal of Philosophy*, Vol. 85, No. 6 (1988), pp. 283-302; M. Baron, 'Impartiality and Friendship', *Ethics* 101 (1991), pp. 836-857.

²⁷ The *locus classicus* is W. Godwin, *An Enquiry Concerning Political Justice*, Book II, Chapter II, pp. 49-54.

²⁸ I borrow these labels from J. McMahan, 'Nationalism and the Demands of Impartiality', in: R. McKim and J. McMahan (eds.), *The Morality of Nationalism* (New York: Oxford University Press, 1997), pp. 107-138, specifically pp. 116-117.

So, it is not inconceivable that the general principle that 'One's Children Come First', or indeed 'One's Country Comes First', could still be endorsed. For all that formal impartiality demands is that partiality in action rest on a principle that is reasonably acceptable to all. We do not always have to treat everyone equally, but if we prioritise some over others, then this should be a kind of partiality that remains within the bounds of impartiality at the higher level of justification. In other words, if they could be shown to derive from our universal duty to humanity, there would still be local obligations to family, friends and the state.²⁹

Assimilationism, then, purports to explain how we could carve out space for special obligations and partiality within impartial morality at the formal level. However, two objections could again be levelled against this attempt. The first objection is much bolder than the second. It claims that the whole idea of a derivation from impartiality is essentially misguided as it makes associative obligation dependent on impartial morality. The second objection is weaker. It merely claims that there is currently no successful derivation of our inherited political obligations from moral duty.

2.1 One Thought Too Many

According to the strong objection, assimilationism will always fail to deliver a plausible justification for special obligation. Briefly, the criticism is that impartial deliberation is both unnecessary and inappropriate in a context that is considered to be deeply personal. So even if a derivation of special obligation from moral duty would succeed, assimilationism would still get the wrong end of the stick. For it would simply fail to capture the intimacy and trust that is implicit in these obligations.

The point is most clearly made in the context of familial obligations. Typical of these obligations is their robustness. Ask any parent why he cares for the well-being of his

²⁹ See J. Rawls, *A Theory of Justice*, pp. 132-133, and for utilitarianism, C. D. Broad, 'Self and Others', D. R. Cheney (ed.), *Broad's Critical Essays in Moral Philosophy* (London: Allen & Unwin, 1971), pp. 262-282, at p. 280.

child and the normal response is 'because it is *my* child'. That relational fact is sufficient reason for most people. But more importantly, there is also something distinctly odd about a parent who would need an external justification. Such justifications would appear to be deeply antithetical to the bond between parent and child.³⁰ Apart from being superfluous, impartiality is also unbecoming of a good parent. Williams famously captures this last point with his phrase that a derivation from impartiality requires us to think 'one thought too many'. He considers the archetypal case of the man who can either rescue his drowning wife or a perfect stranger. While assimilationism agrees that the man is right to rescue his wife, it involves

the idea that moral principle can legitimate his preference, yielding the conclusion that in situations of this kind it is at least all right (morally permissible) to save one's wife... But this construction provides the agent with one thought too many: it might have been hoped for (for instance, by his wife) that his motivating thought, fully spelled out, would be the thought that it was his wife, not that it was his wife and that in situations of this kind it is permissible to save one's wife.³¹

Williams's objection is that a husband should come to the aid of his wife, not for the additional thought that a universal principle dictates that he should, but specifically and only 'because she is *his* wife'. Assimilationism evokes the uncomfortable image of a husband coolly calculating the moral benefits and costs of his action, while a more full-blooded and unreserved expression of loyalty would be more appropriate.³²

A similar view underpins Oldenquist's account of patriotism. The impartialist attempt to acclaim a form of patriotism, for example, fails on the following ground:

An impartial patriot says that he supports his country just because it has features *a . . . n*, and therefore he is committed to the position that if his country and its adversary exchanged the features *a . . . n*, he should support the adversary. If it is true that he supports his country solely because he thinks it is in the right, it follows that the fact that one of the countries is his does not enter at all into his decision. Such 'patriots' believe themselves to be objective, like

³⁰ See also J. Horton, *Political Obligation*, pp. 147-148.

³¹ B. Williams, 'Persons, Character, and Morality', pp. 214-215.

³² For this point, see also A. Oldenquist, 'Loyalties', *The Journal of Philosophy*, Vol. LXXIX, No. 4 (1984), pp. 173-193.

ideal observers, and simply lucky that their own country is in the right, whereas I think that most of us believe that their underlying motives usually are pure loyalty and their reasons rationalizations.³³

Oldenquist taps into the same intuition that the judgment of the true patriot 'is partly determined by the fact that the good of *his* country is at stake.'³⁴ The patriot who first needs the assurance that his country is in the right – like the husband or the parent who first needs to flip a coin, so to speak, and then proceeds to rescue his wife or daughter – stands accused of having no loyalty at all, precisely because he requires impartial grounds. The key-problem, of course, is that the coin *could* have fallen the other way up, even if in fact it did not; impartial reason could just as well have gone the other way, which is why we think assimilationism makes for lousy patriots, husbands and parents. For we expect from friends and mothers and patriots, at least sometimes – though not necessarily always – that they will help us even if that were *not* justified by impartial principles. In other words, assimilationism makes associative obligations entirely conditional on higher-order, impartial morality, while our understanding of these obligations suggests that the two may conflict in fundamental ways.

There is I think a great deal of truth in this objection. Nevertheless, there are also some persistent doubts. Many critics have noted how much the associative argument relies on the analogy between the family and the state.³⁵ The idea that assimilationism is inappropriate may seem incontrovertible in the example of our family-members and those relations that are relatively intimate and close. But the idea is not so obvious in the case of large-scale associations like states. What is more important, there is the further problem that our understandings of political and associative obligations also suggest that they are not quite so unconditional at all. The problem, after all, is that states, like

³³ Ibid., p. 183.

³⁴ Ibid., p. 184.

³⁵ Most influentially, J. Locke in his *Second Treatises of Government*. For recent examples, see A. J. Simmons, 'Associative Political Obligations', *Ethics*, 106 (January 1996), pp. 247-273; C. Heath Wellman, 'Friends, Compatriots, and Special Political Obligations', *Political Theory*, Vol. 29, No. 2 (April 2001), pp. 217-236; R. Dagger, 'Membership, Fair Play, and Political Obligation', *Political Studies*, 2000, Vol. 48, pp. 104-117.

families, are often malicious or even criminal. As impartialist critics delight to point out, do we really want to say that there are associative obligations in these cases?

2.2 Justice, Expedience and Political Obligation

On balance, the strong objection may therefore be less than conclusive. To see why the assimilationist model nevertheless fails, it is necessary to turn to the weaker objection. As we saw, this objection claims that currently there is no derivation of moral duty that actually succeeds in grounding political obligations. While this objection does not reject a derivation from higher-level impartiality on principle, it argues that all attempts to do so have thus far failed.

To expose this failure, we have to turn to the view, shared by theorists such as Kant, Rawls, Waldron and Goodin, that political obligation is the corollary of our natural duty of justice.³⁶ According to this view, political institutions are a prerequisite to justice. Briefly, the first premise of the argument is that justice is a natural duty. The second premise is that justice can never be achieved without political organisation and positive law. It then follows that individuals have a natural duty to establish, concur with and support political institutions, provided that these institutions do indeed promote justice.

This explanation faces the following challenge. The special allegiance objection points out that political obligations are member-specific. However, it is not immediately clear how this scope-restrictedness could be accommodated within a natural duty framework. The natural duty to advance justice obviously binds universally and

³⁶ Kant speaks of the civil state not as a union for some common end that the contractors all agree to share, but 'as an end in itself which they all *ought to share*.' It is our natural duty to submit to civil authority and positive law because justice can only be attained within a civil state and the rule of law. See his 'On the Common Saying: 'This May be True in Theory, but it does not Apply in Practice', in: H. Reiss (ed.), *Kant: Political Writings* (Cambridge: Cambridge University Press, 1991), pp. 61-92, especially p. 73. For a more recent statement of the principle, see J. Rawls, *A Theory of Justice*, pp. 114-117, 334, where Rawls subscribes to the view that justice is a natural duty, and goes on to claim (p. 115) that 'This duty requires us to support and to comply with just institutions that exist and apply to us.' In a similar frame of mind, J. Waldron, 'Special Ties and Natural Duties', *Philosophy and Public Affairs*, Volume 22, Issue 1 (1993), pp. 3-30, at p. 27, writes that an 'organization that is just, effective and legitimate (in the sense of being singled out as *the* salient organization for this territory) has *eo ipso* a claim on our allegiance.'

generally. But if associative obligations are derived from the natural duty to facilitate justice, then why do we not owe such obligations to all just institutions and states?

Waldron suggests in response that the principles of justice are 'range limited'. By this he means that these principles distinguish between 'insiders', who are to secure just institutions among themselves, and 'outsiders', whose interests are not affected by what goes on in the particular institutional context to which justice applies. Whereas the insiders of just institutions continue to have the full associative obligation to submit to the authority of the law or, at any rate, to divulge the right kind of loyalty to the state, outsiders are merely under the weaker duty not to interfere or obstruct the course of justice.

However, it is far from clear that justice actually lends support to the insider-outsider distinction. 'Substantively,' Waldron writes, 'an individual is within the range of a principle if it is part of the point and justification of the principle to deal with his conduct, claims, and interests along with those of any other persons it deals with.'³⁷ But how can we justify making such a distinction? Citing Kant, Waldron then tells us the following little story. In the state of nature we all have a duty to found political associations such that we may enter in peaceful, stable and just relations. However, it would be quite pointless to immediately seek to secure just institutions for each and everyone. On the whole, it is far more expedient to begin by establishing a just order among those people living in one's immediate proximity, and with whom conflict over the appropriation of resources is more likely to occur. Thus, political regimes enter into the world on a local scale and their regulatory force extends only to a bounded territory, and those residing in it. Drawing boundaries, in short, is simply the most efficient way to accomplish a just world.

But the appeal to expedience in this story is ultimately too thin. It may well be right to circumscribe the scope of justice to certain insiders, but it is of course highly unlikely

³⁷ J. Waldron, 'Special Ties and Natural Duties', p. 13.

that these restrictions will conveniently coincide with the territorial boundaries of the state. I can see, perhaps, why Waldron's story makes sense in a state-of-nature-scenario, when justice and the rule of law are just starting to emerge from primordial madness, but in more advanced scenario's further considerations come in view. For example, Waldron acknowledges that the 'poorest person in New Zealand is considerably better off than most people in Bangladesh'.³⁸ But if justice includes the fair distribution of basic resources, why should it be conducive to that ideal to limit a New Zealander's duty to secure justice to New Zealand?³⁹

Robert Goodin has attempted to distinguish insiders from outsiders in a similar but perhaps more sophisticated manner.⁴⁰ His attempt is grounded in the consequentialist duty to promote well-being across the board. While this duty is general, his thesis similarly claims that it is performed more efficiently within truncated structures that differentiate between members and non-members. In his view, special duties are merely 'distributed general duties'.

That is to say, special duties are in my view merely devices whereby the moral community's general duties get assigned to particular agents... A great many general duties point to tasks that, for one reason or another, are pursued more effectively if they are subdivided and particular people are assigned special responsibilities for particular portions of the task.⁴¹

³⁸ Ibid., p. 14.

³⁹ In *On Nationality*, p. 63, David Miller rightly protests: 'Why does it make sense to assign responsibility for the rights and welfare of Swedes to other Swedes and the rights and welfare of Somalians to other Somalians, if we are looking at the question from a global perspective?' Scheffler similarly argues that if we take distributive justice seriously, we seem stuck with 'the idea that associative duties provide a mandate for those who are already rich in resources to turn their attentions inward, and largely to ignore suffering and deprivation in the rest of the world'. See his 'Families, Nations, and Strangers', at p. 58. For his 'distributive objection', see also the essays 'Liberalism, Nationalism, and Egalitarianism', pp. 66-81, and 'The Conflict between Justice and Responsibility', pp. 82-96, all reprinted in *Boundaries and Allegiances*.

⁴⁰ R. E. Goodin, *Protecting the Vulnerable* (Chicago: University of Chicago Press, 1985). See also R. E. Goodin, 'What Is So Special about Our Fellow Countrymen?', *Ethics*, 98 (July 1989), pp. 663-686; R. E. Goodin, 'The State as Moral Agent', in: P. Pettit & A. Hamlin, (eds.) *The Good Polity: Normative Analysis of the State* (Oxford: Basil Blackwell, 1989), pp. 123-139; R. E. Goodin and P. Pettit, 'The Possibility of Special Duties', *Canadian Journal of Philosophy*, 16 (December 1986), pp. 651-676.

⁴¹ R. E. Goodin, 'What Is So Special about our Fellow Countrymen?', pp. 678-679.

In other words, these special obligations result from a division of moral labour. Goodin furnishes the following example. Drowning swimmers are entitled to assistance on the basis of the general duty of beneficence. However, this does not imply that we should all immediately rush into the water. It is far more efficient to assign this responsibility to a handful of lifeguards, who are selected on their ability to carry out such operations. Other examples come to mind. While all children are generally entitled to care, the responsibility to discharge this general duty is best assigned to the parents. And we can now also see how political obligations could emerge. All people have a claim to assistance in need. But again, this does not necessarily mean that we ought to provide for the needy all over the world. It is probably more efficient to assign this task primarily to states operating within a confined territory.

But we still need to know how such distributions of special obligations come about. How do we know, say, that I have a special duty to Johnny and not to Sonny? Goodin's answer is that special duty is a function of a constant factor, which is the general duty of beneficence, and a variable x that accounts for the allocation of special tasks to specific people. This variable is to be understood as a principle of responsibility, which must be read in consequentialist fashion. So the principle of responsibility directs our general duty to those persons who are particularly vulnerable or susceptible to our actions or omissions. We all know that some people are more vulnerable to what we do (or fail to do) than others. If my actions or omissions cannot possibly benefit or harm Sonny, I have no special duties to him. There is no relation of vulnerability between Sonny and myself. If my actions or omissions have an effect on Donny's well-being, then I acquire a special obligation to Donny. However, if my actions have a greater impact on the well-being of Johnny, then my obligation to Johnny is even stronger. In this way the strength and urgency of our duties begins to vary.

Applying this logic, Goodin insists that his model allocates duties roughly in accordance with commonsense morality.⁴² Our own well-being, the well-being of our children, neighbours, and so forth, are obviously more dependent on our decisions and actions than the well-being of distant others, or so he claims. For that same reason, citizens have an obligation to primarily promote each other's well-being rather than the well-being of humanity at large. As he puts it,

... persons relatively near to us in space and in time probably *will* be rather more vulnerable to us... This fact saves my argument from the traditional *reductio* of requiring that we give everything we have to starving Asians or that we forever save everything for infinitely receding future generations...⁴³

Furthermore, by appealing to vulnerability, political associations and institutions can be established to facilitate the administration of certain of these special tasks and duties.

Now, the idea that we live in an institutionalised web of social relations from which we derive non-voluntary moral obligations is, in my view, convincing enough. But the main problem with Goodin's model is that its distributive criterion, vulnerability, is not entirely independent of the duties it is supposed to distribute in the first place. Hence, the argument succumbs to a vicious kind of circularity. Take the example of parental duties. Superficially it may be true that children are most vulnerable to their natural parents. However, going one step further, one could wonder why this is so. One plausible answer must surely be that this vulnerability arises precisely because convention assigns the duty to rear, feed and educate children to their natural parents. Given this conventional model of parental duty, children are indeed specifically vulnerable to parental neglect because there is no one else who feels responsible in the same deep way. But what if convention assigned that responsibility to grandparents or, say, in Platonic fashion, to some state-nursery run by philosophers? It might also be correct to say that in welfare states people are vulnerable to their fellow-citizens and the state. But having said that, is this vulnerability not due to the fact that we have welfare

⁴² R. E. Goodin and P. Pettit, 'The Possibility of Special Duties', p. 666.

⁴³ R. E. Goodin, *Protecting the Vulnerable*, p. 121.

states in the first place? Given the fact that we have awarded so much responsibility to the state, it should hardly be surprising to see that we have become more vulnerable to our fellow-citizens, and less to friends, relatives and local communities.

Ultimately, then, vulnerability is a problematic distributive criterion of duty because the vulnerability between persons is co-determined by how duties are *de facto* assigned by traditional conventions and historical patterns. These institutions allocate duties that in their turn distribute vulnerability rather than the other way around. Goodin anticipates this problem, but he does not seek to deny it.⁴⁴ The allocation of duty, he says, is indeed a mere social fact, but he argues that vulnerability is nevertheless required for the moral justification of that allocation. So, while he accepts that special responsibilities are assigned to people simply on the basis of historical patterns, he also insists that the mere fact of such institutions is insufficient to make these obligations morally binding. It is only the element of vulnerability that is fostered by these patterns that allows us to say that special obligations and associative loyalties are normative: 'The pragmatic moral 'ought' ('What ought I to do?') attaches to the vulnerabilities, rather than to conventionally defined duties and responsibilities *per se*.'⁴⁵

Formally this might be a tenable position. Strictly speaking, we do not need to espouse what Goodin describes as the 'traditionalist cum intuitionist' position regarding your station and its duties. We can maintain that we have special duties, not because social convention determines that we do, but only because convention places people in a vulnerable position to us. We can argue, indeed, that it is only with the added element of vulnerability that people enter into moral relations. But unless we say something about the desirability of these *de facto* allocations of vulnerability, it does not appear to be much more than a formal point. Does Goodin's thoroughly theorised model really add anything to, as he scathingly calls it, the 'most primitive form of moral consciousness'? Goodin says that parents have special duties to their children because our institutions

⁴⁴ *Ibid.*, pp. 124-125.

⁴⁵ *Ibid.*, p. 125.

are such that the latter are specifically vulnerable to the former. Primitivism says that we acquire these duties simply because they inhere in the institution of parenthood. And to be frank, I do not see all that much difference.

We may agree, then, with the general point behind this model that some division of labour between state-like institutions is likely to increase the sum-total of human well-being. However, it is a big leap from this insight to the justification of our political obligations to specific states as we have inherited them from our ancestors. Goodin's approach simply fails to cope with the contingent nature of human vulnerabilities. Moreover, it is simply false to believe that our family members or fellow-countrymen are always more vulnerable than the next person. To put it differently, while in some circumstances it may be perfectly permissible for husbands to rescue their wives rather than a stranger, other circumstances may necessitate a different course of action. For example, there may be a far better chance of successfully rescuing the stranger's life. Why still rescue your wife? The problem, in short, is that there does not appear to be an easy and internally stable rule that explains how special duties are appropriately distributed in accordance with the consequentialist duty to human-kind.

3. Associativism: One Thought too Few?

The impartialist's toolkit contains explanations for how the constituency of moral duty could be reduced from the universal to a local constituency. However, for the reasons I have just set out, these tools are simply too blunt to craft an adequate account of our inherited political obligations.⁴⁶ This failure leaves liberal impartiality exposed to

⁴⁶ C. Heath Wellman, in his 'Friends, Compatriots, and Special Political Obligations', *Political Theory*, Vol. 29, No. 2 (April 2001), pp. 217-236, at pp. 231-232, agrees that assimilationism has not been very successful. Yet he is convinced that its problems are merely due to 'the unfortunate unwillingness' of contemporary liberals to confront this issue. However, if, in recent years, political theory has mainly occupied itself with 'more fashionable concerns like the difference principle', then this is perhaps because such a justification is precisely *not* forthcoming.

radicalism. In the remainder of this chapter, I want to develop the associativist alternative.

As a general thesis, associativism asserts that relational facts possess obligation-creating significance. To put it briefly, associativism maintains that statements such as ‘because she is my wife’, ‘because he is my friend’, and so forth, need no external justification. The model claims something similar for political obligation. We have such obligations purely because we are members of a particular state, without needing a derivation from moral duty.

Now, we already saw that political obligations are usually inherited. We are obligated as a result of the fact that previous generations built particular political structures, constitutions and laws. How does associativism explain this? Briefly, associativism claims that we have to reject the Jeffersonian premise that inevitably steers us towards consent and natural duty theories. That is, we need to jettison the view that generations are necessarily like sovereign nations. Instead of this view, we need to conceive of the possibility that states could be transgenerational associations. More precisely, we need to claim that rather than each generation forming a separate agent, generations could be related in such a way that they form a corporate person. On this view of corporate personality, generations are linked in the manner of the temporal stages of one continuous person. This means that they are related such that the collective intention of earlier generations continues to be authoritative over the present.

But the idea that relational facts are primitive in this way provokes the objection that associativism ‘thinks one thought too few’. In fact, there are at least two problems that stand out. First, we still need to explain *why* generational relations should have this basic obligation-creating significance. This explanation cannot be deduced from universal duty, but some clarification is still in order. Second, if it is claimed that a relational fact has *basic* obligation-creating status, then this suggests that *all* similar facts must have this status. In other words, we still have deal with the pivotal objection

that transgenerational associations may be criminal and corrupt. I shall discuss the last objection in Section 4 of this chapter. For now, however, I want to focus solely on the first problem.

3.1 *The Conceptual Argument*

Associativism is often claimed to derive support from the notion that being a member is already conceptually linked to being the subject of associative obligations.⁴⁷ In other words, the contention is that our understandings of certain relationships and associations are normatively laden. As we have already seen, we use the term friendship to describe a particular kind of relationships that is based on loyalty, mutual respect as equals and trust. Thus, if you were to treat me in a purely instrumental way, you could not qualify as a friend. Friendship simply is the equivalent of acknowledging certain mutual obligations. On this view, too, compliance to the law is similarly inherent in the meaning of concepts like citizenship, authority and the state. By rejecting the legitimacy of your political obligation, you are rejecting citizenship, law and the state.

However, revealing such conceptual links is one thing, justifying political obligations quite another. The critics of this view have rightly complained that this argument obfuscates the real issue at hand.⁴⁸ For while these links exist, we still need to show why such obligations are normative. And for this we need to clarify why concepts like citizenship, law and the state, together with their implicit normative cargo, should appeal to us in the first place. So, it may be correct to note that our understandings of citizenship, law and the state are, as Dworkin puts it, 'pregnant' with obligation. But it is not valid to conclude from this that the issue of justification has thus been solved.⁴⁹ That conclusion would merely succeed in moving the problem from one place to another. Expounding the normative connotations hidden in our conceptual

⁴⁷ For a discussion of this position, see J. Horton, *Political Obligation*, pp. 137-145.

⁴⁸ For example A. J. Simmons, *Moral Principles and Political Obligation*, pp. 38-43.

⁴⁹ R. Dworkin, *Law's Empire*, p. 206.

understandings of relations does not suffice to show why we should want these associations.

3.2 Identity-dependent Arguments

The most popular argument for associativism appeals to people's identification with groups and associations. On this view, special obligations are grounded in communal attachments that are constitutive of our sense of selfhood. The fact that one identifies with an association, on this view, is a necessary and sufficient condition for special obligations to emerge. To put it differently, this approach agrees that many of our relationships are conceptually pregnant with obligations. But for the additional normative argument, it turns to the fact that people define their identities in terms of the associations to which they belong. So, the argument is *identity-dependent*. According to Hardimon, the gist is that (1) the community we identify with is not just a figment of our own imagination, (2) mutual obligations are inherent in our understanding of the community, and (3) when people identify with a community, they view themselves in such a way that it simply becomes illogical to resist institutional norms and obligations of roles and membership.⁵⁰ Concurring with this last point, De-Shalit writes:

... if and when one admits the existence of a community, and if one acknowledges that the community constitutes one's identity, then it is absurd at the same time to deny any obligation to the community and its members. If one acknowledges the importance of the community, then one wishes the community to be sustained, and even to flourish.⁵¹

Miller accounts for the obligations between fellow-nationals on the same basis:

⁵⁰ M. O. Hardimon, 'Role Obligations', *The Journal of Philosophy*, 91, Number 7 (1994), pp. 333-363, at p. 358.

⁵¹ A. De Shalit, *Why Posterity Matters*, p. 15. Horton writes in his *Political Obligation*, p. 157: 'It should not be surprising [...] that some institutional obligations, through their deep-rooted connections with our sense of who we are and our place in the world, have a particularly fundamental role in our moral being. That these kinds of institutional involvement generate moral obligations, and that these obligations rather than standing in need of justification may themselves be justificatory, is only to be expected.' See also the discussion in Y. Tamir, *Liberal Nationalism* (Princeton: Princeton University Press, 1993), pp. 95-116.

Because I identify with my family, my college, or my local community, I properly acknowledge obligations to members of these groups that are distinct from the obligations I owe to people generally. Seeing myself as member, I feel a loyalty to the group, and this expresses itself, among other things, in my giving special weight to the interests of fellow-members.⁵²

While this view has certain merits, it also suffers from the weakness that it does not distinguish between a sense of duty and its rational justification. It may well be that constitutive attachments explain why members want their communities to flourish, or why they are more concerned about the well-being of fellow-members than non-members. Indeed, it probably is unusual and perhaps even absurd for someone to identify with a community and yet to completely ignore its interests at the same time. But does this intimation of a duty mean that we actually have such obligations? One problem is that some people altogether fail to identify with their country to begin with. Another problem is that, as a foundation for associative obligation, self-identification seems rather indiscriminate. We identify with many groups. Aside from our country, we feel loyalty to the firm we work for, a local football club and what have you. I may derive a *sense* of obligation from these attachments, but why should this be sufficient reason to bind me morally?⁵³

3.3 Value-dependent Arguments

In an essay on nationalism, Simon Caney has drawn attention to a range of ‘value-dependent arguments’ that he thinks are more appropriate for the job.⁵⁴ The general idea is that associativism is justified if the association or relationship in question is valuable

⁵² D. Miller, *On Nationality*, p. 65.

⁵³ For examples, see C. Heath Wellman, ‘Relational Facts in Liberal Political Theory’, p. 555, ff.; A. J. Simmons, ‘Associative Political Obligations’, pp. 261-265; J. McMahan, ‘Nationalism and the Demands of Impartiality’.

⁵⁴ S. Caney, ‘Individuals, Nations and Obligations’, in: S. Caney, D. George, P. Jones (eds.), *National Rights, International Obligations*, (Oxford: Westview Press, 1996), pp. 119-138, at pp. 130-133. Caney does not believe this argument is successful in the case of nationality. However, his general assessment of value-dependent arguments seems to be more optimistic. See also S. Scheffler in ‘Relationships and Responsibilities’, in: *Boundaries and Allegiances*, pp. 97-110.

in itself. Friendship is a plausible example.⁵⁵ Obligations of friendship are justified because (1) obligations are intrinsic in friendship and (2) friendship is something intrinsically valuable. However, Caney does not think that such an argument can be successfully mounted in the case of our special obligations to fellow-nationals. The idea that nationhood is intrinsically valuable like friendship, he claims, is far too contentious. So, the question is how value-dependent arguments could help to explain our obligations to a political community.

Andrew Mason has furnished another example of a value-dependent argument.⁵⁶ According to Mason, the justification of our political obligations ultimately originates in the value of citizenship rather than in the value of nationality.

Citizenship has intrinsic value because in virtue of being a citizen a person is a member of a collective body in which they enjoy equal status with its other members and are thereby provided with recognition. This collective body exercises significant control over its member's conditions of existence (a degree of control which none of its members individually possesses). It offers them the opportunity to contribute to the cultural environment in which its laws and policies are determined, and opportunities to participate directly and indirectly in the formation of these laws and policies. Part of what it is to be a citizen is to incur special obligations: these obligations give content to what it is to be committed or loyal to fellow citizens and are justified by the good of the wider relationship to which they contribute.⁵⁷

Because (1) political obligations are conceptually implicit in the idea of citizenship and (2) citizenship is an intrinsic good, associative obligations are truly normative. It is the latter part of the argument that does the justificatory work in Mason's account.

But his argument also has certain weaknesses. One potential weakness is the republican conception of citizenship as an intrinsic good.⁵⁸ By describing citizenship as an intrinsic good, Mason suggests that citizenship is not just valuable because it is instrumental to another good, but also because it is part of the good in itself.

⁵⁵ Here, Caney builds on J. Raz, 'Liberating Duties', *Law and Philosophy*, Vol. 8 (1989), pp. 3-21.

⁵⁶ A. Mason, 'Special Obligations to Compatriots', *Ethics*, 107 (April 1997), pp. 427-447.

⁵⁷ *Ibid.*, p. 442.

⁵⁸ For a critical view, see C. Heath Wellman, 'Friends, Compatriots, and Special Political Obligations', pp. 222-224.

Citizenship, he says, is 'a constituent of well-being' rather than 'a means to well-being'.⁵⁹ However, the argument designed to explain why this is the case is not completely convincing. Mason argues that citizenship provides us with social recognition as equals. However, I doubt that that citizenship in modern states can really satisfy the need for social recognition. Representative democracies, after all, are not like ancient republics. In addition, Mason claims that citizenship gives us greater leverage and control over our cultural environment, public policy and law. However, unless we assume that leverage and power is an intrinsic good itself, this claim remains instrumental. What is more, in this form, the argument does not account for the fact that political obligation is typically passed on from one generation to the next. Mason does not explain why political association and obligation is a relationship between generations.

Despite these problems, the value-dependent strategy provides what is in my view the most attractive account of associativism. I want to employ this theory by building on the notion of corporate personality and integrity as set out in previous chapters. My argument is value-dependent, but it appeals to the intrinsic value of integrity rather than that of nationality or citizenship. In this form, I argue, the theory avoids some of the problems linked with Mason's account. Essentially, the claim is that we inherit certain political obligations because (1) to have personal identity is to recognise the normativity of earlier intentions and (2) personal identity is of intrinsic value.

To begin with, consider what these two claims imply in the instance of individual persons. First, the claim that earlier selves could bind later ones is captured by the idea of integrity. Integrity means that the intentions and commitments of earlier selves have normative weight for us now. We cannot just disown our past actions. Integrity is opposed by the notion of a time-slice agent, which prescribes a mode of thought in which past intentions and actions are like external facts. Second, the argument that

⁵⁹ A. Mason, 'Special Obligations to Compatriots', p. 441.

integrity has intrinsic value rests squarely on the practice theory of personal identity. According to this theory, whether we should be temporally extended or time-sliced agents, moderates or extremists, is itself a normative question. One branch of the practice theory claims that integrity is integral to the idea of the goodness of a life as a whole. Integrity is what makes us into a unified person, capable of having a unified good. A second argument is that integrity facilitates responsibility.

If we want to make the same case for political obligation, we have to apply the notion of integrity at the wider level of the state. First, the claim that a present generation of citizens is bound by earlier generations is summed up in the model of corporate personality. The idea is that the collective intentions and the commitments of earlier generations continue to have normative weight for present and future generations. We cannot just disown the actions of past generations. This model of collective integrity stands in sharp contrast to the Jeffersonian model of generations as independent agents. The latter model regards future-directed intentions and commitments, as we find them in founding constitutions and treaties, as tyrannical. Both models are normative in the sense that they advocate different views about what our attitude to the collective past should be.

Second, we again need to appeal to the practice theory of personal identity to explain why corporate integrity, too, is a virtue. One important claim, which I advanced in Chapter VI, continues the argument from the goodness of a life as a whole. The claim is here that present generations participate in shared projects and goods, whose achievement requires co-ordination and co-operation between several generations. It is for the sake of this transgenerational nature of the human good, so the argument goes, that generations should be normatively connected to each other. A second approach, which I have left largely unexplored until now, continues the argument from responsibility and action ownership. It points out that corporate personality explains why later generations could be held responsible for the actions of earlier ones. Thus, it

facilitates a view on which accountability does not just slip through the cracks of time, but passes on from one generation to the next.

There might be other normative arguments that underscore the importance of integrity. But whatever these arguments are, the point is here that the value-dependent strategy explains why associativism, as an account of political obligation, is a plausible theory. Inherent in corporate personality is precisely the view that one generation could bind other generations, share responsibilities and interests, as it is inherent in the notion of one individual person that one person-stage could bind another, and share responsibilities and interests. By showing that the relation of integrity facilitates transgenerational goods and transgenerational responsibility, we have an account of why associative obligations are normative.

4. Bad Associations

Rather than appealing to self-identification or the intrinsic value of nationhood or citizenship, I have suggested that political obligations are immanent in the idea of corporate personality, and that integrity is a virtue not derivable from a more basic virtue. However, critics often object that associativism fails to appreciate that relational facts greatly vary in character, substance and purpose. Heath Wellman points out:

... if there is truly magic in the pronoun 'my', then we apparently must have special duties to all our relations, just as the magic in promising means that we have special moral duties to keep all of our promises. Such a view is untenable, however, because it is absurd to posit special duties to all relations.⁶⁰

Some relations, like some promises, are morally more attractive than others, but associativism makes us blind to this fact, or so the allegation goes.

To demonstrate this point, the critics begin by citing numerous examples of associations that foster unjust practices or immoral ends. Dagger notes: 'Philanthropic

⁶⁰ C. Heath Wellman, 'Relational Facts in Liberal Political Theory', p. 553.

groups have members, but so does the Mafia.⁶¹ Heath Wellman points to racist and sexist groups. And according to Simmons, 'local practices and institutions can be unjust, oppressive, pointless, woefully inefficient, and in other ways normatively defective.'⁶² They then ask whether associates of a defective association, say, the Mafia, are bound by the organisation's codes of silence and loyalty. Finally, because the Mafia is a relational fact, it is pointed out that associativism is compelled to defend the opposite view. Since it claims that such facts have obligation-generating significance, associativism must also hold – absurdly – that Mafia membership generates obligations. Since this is assumed to be self-evidently mistaken, associativism must clearly be wrong. For we are only able to nullify associative obligations to the Mafia once we move to assimilationism, and subject these obligations to review from the viewpoint of impartial duty. To sum up, we need to bring in principles of justice to either certify or declare void special or political obligations.

Usually, associativists respond to this challenge in one of the following two ways. First, some associativists bank on including a proviso, effectively claiming that associative obligations are only conditionally valid. That condition invariably states that the practices sustained by the association, its character, substance and purpose, should be consistent with an external norm or a principle of justice. For example, according to Hardimon, 'it is crucial to distinguish between the idea that one might be under a moral obligation simply because one occupies a particular role *given that certain moral background conditions are in place* (for example, that the institution is just and that the role is reflectively acceptable) and the idea that a moral role could generate moral obligations *regardless of the moral background conditions*.'⁶³ But including such a robust proviso hollows out associativism to an extent that it leaves unclear why we need

⁶¹ R. Dagger, 'Membership, Fair Play, and Political Obligation', p. 110.

⁶² A. J. Simmons, 'Associative Political Obligations', p. 266. For what he describes as the 'obnoxious identity objection', see also S. Caney, 'Individuals, Nations and Obligations',

⁶³ M. O. Hardimon, 'Role Obligations', p. 351.

this thesis at all.⁶⁴ In the end, Hardimon is not really an associativist at all. Certainly, he rejects voluntarist approaches, but because he insists that role obligations must be 'reflectively acceptable' his approach slips back into what is essentially a hypothetical mode of contractualism reminiscent of the natural justice models advanced by Kant, Waldron and Goodin.

A similar relapse occurs in Dworkin's treatment of the issue. However, here the caveats are largely smuggled in through the backdoor. In *Law's Empire*, Dworkin wants to distinguish between 'bare' and 'true' (or 'fraternal') communities.⁶⁵ A bare community refers to a loose grouping of individuals who are physically brought together by biological, geographical and historical factors. The existence of a bare community is a precondition for a fraternal community to emerge. However, a community is 'fraternal' only if it meets a further number of conditions. First, its members must view their associative obligations as *special*. Second, these obligations must also be seen as *personal*, meaning that I cannot be released from my duties to a specific fellow-member merely by benefiting the group at large. Third, these obligations must be taken to express *concern* for the well-being of the members. Fourth, and most notably, there must be an *equal* concern for all members. 'True' communities, Dworkin claims, are 'conceptually egalitarian'. Obviously, this last condition is intended as a check on associations that harbour basic inequalities, like caste societies. According to Dworkin's classification, such societies do not even count as associations. Moreover, he adds that if certain morally repugnant associations should slip through this maze of caveats, then associative obligations are simply cancelled by 'severe and deep' injustices.⁶⁶

⁶⁴ Here, I agree with Tamir's claim that 'If only morally valuable communities could generate associative obligations, the latter would become a meaningless concept.' See *Liberal Nationalism*, p. 102. In his 'Relational Facts and Liberal Political Theory' (p. 555), Wellman refers to the proponents of this view as 'closet assimilationists'.

⁶⁵ R. Dworkin, *Law's Empire*, p. 201.

⁶⁶ *Ibid.*, pp. 202-206. For a dissenting voice, see J. Horton, *Political Obligation*, pp. 156-157.

The second and in my view more credible solution suggests that we bite the bullet. On this view, associative obligations persist even in the face of severe and deep injustice. As Tamir argues,

There is no reason to assume, as Dworkin does, that only membership in morally worthy associations can generate associative obligations. For example, members of the Mafia are bound by associative obligations to their fellow members, meaning that they have an obligation to attend to each other's needs, to protect each other, to support the families of those killed 'in action', and the like.⁶⁷

But, as she points out, this view is actually less far-fetched than most assimilationists assume it to be. Clearly, associativism endorses such obligations only on a *prima facie* basis. While associative obligation may be one factor in my judgment about what I ought to do *all things considered*, this is not the only consideration. So, associativists do not deny that there are other, moral obligations and concerns that could overrule what we owe in virtue of our membership to a particular association.

This solution, too, is claimed to have drawbacks. First, we have to accept a form of moral pluralism that claims that neither impartiality nor associativism is capable of going it alone.⁶⁸ At the level of ultimate judgements, the potential for a deep and all too familiar conflict remains. Second, even at the *prima facie* level, assimilationists will maintain that there is something absurd about the suggestion that we could have an obligation to unjust or criminal associations. Assimilationism appeals to the intuition that the members of such associations are under no such obligation at all. If Tony instructs Donny to kill Sonny, it is erroneous to think that Donny acquires a *prima facie* obligation to assassinate Sonny, which is only later overruled by a general, moral obligation. There can be no obligation to murder Sonny to begin with.

⁶⁷ Y. Tamir, *Liberal Nationalism*, p. 101. For a similar argument, see also A. Oldenquist, 'Loyalties'; and M. Gilbert, 'Agreements, Coercion, and Obligation'.

⁶⁸ For example, Oldenquist claims: 'In terms of the logic of reasons they provide, loyalties are a third category of the normative, distinct from both self-interest and impersonal morality.' See his 'Loyalties', p. 176.

What could we say in response to these objections? It is important to begin by pointing out that although the problem of bad associations may not entirely disappear, the value-dependent underpinnings of associativism offer more room for manoeuvring than identity-dependent accounts. If, indeed, identification is a sufficient and necessary condition, then we have to agree with Tamir that moral provisos and caveats are, like a ghost from the machine, introduced all too conveniently and with the sole purpose of ruling out the existence of prima facie obligations when someone happens to develop constitutive attachments to associations like the Mafia. Such an insurance policy is inconsistent with the identity-dependent basis of associative obligations itself. However, if we turn to value-dependent accounts, the inconsistency vanishes. Since associative obligations are grounded in the *value* of the association in question, that value already serves as a screen itself. For example, Mason notes that ‘there are limits to what a citizen body can do to their individual members without destroying the relationship of citizenship.’⁶⁹ In his view, states that undermine the political equality of their members also destroy the very grounds on which they can demand loyalty and support. Corporate personality imposes a similar screen. It points out that associative obligations are conditional on the state’s integrity. And there are also limits to what a political community can do without compromising its integrity.

The exact scope of the limits thus drawn by this approach naturally depends on how the content of integrity is fully spelt out. But it is clearly not inconsistent to claim that when a state jeopardises its own integrity, it releases its citizens from their political obligations. So, this value-dependent approach seems more suitable than Dworkin’s approach. Dworkin’s mistake, in my view, is that he does not appeal directly to the value of integrity itself. Instead, he complicates things by putting too much weight on the distinction between ‘true’ and ‘bare’ communities. States need to have integrity to qualify as ‘true’ communities. But it seems that in order to generate political

⁶⁹ A. Mason, ‘Special Obligations to Compatriots’, p. 444.

obligations, states need to possess a mix of many other characteristics as well. And to round things off, he offers an external proviso. Instead, it seems more logical to start with a value-dependent account of what integrity is and why it is valuable. This account imposes its own internal limitations on what a state could do without destroying the foundation of political obligation. States need to respect the laws and principles, the 'intention', embodied in their constitutions. These laws and principles thus function as a filter on our obligations and responsibilities.

It must be conceded that this value-dependent approach may not entirely solve the problem of bad associations. States can do many awful things, or so I believe, and still maintain their integrity. I have argued that integrity is a virtue or value because that is what it takes for both individuals as well as states to qualify as one continuous person, capable of having a long-term good and owing up to action. But having this virtue is not inconsistent with performing a morally defective or wrong action. Two things have to be kept apart here. While it is one thing to say that he or she is a *unified* person, it is another thing to say that he or she is a *moral* person. These judgements invoke a different meaning of the term 'badness'. Korsgaard has argued that when we say that an association or a person or his action is bad, we can mean two things.⁷⁰ First, we can mean in the conventional manner that this person or his action is morally wrong or unjust. This involves an *external* judgment. But second, we can also mean to censure an act in the sense that we say 'This is a bad car' or 'This is a bad free kick'. We could describe this as an internal judgment. My judgment presupposes that there is a standard *internal* to the thing or activity in question, usually expressed in terms of a function or a purpose. A car is usually meant for driving and the purpose of a free kick is to score a goal. From the internal standpoint, a bad car fails 'as a car'. It fails to live up to what it means to be a good car. And a bad free kick fails 'as a free kick'. It fails to live up to what free kicks aim to achieve.

⁷⁰ C. M. Korsgaard, 'Self-Constitution in the Ethics of Plato and Kant', pp. 12-15, and more generally *The Sources of Normativity*.

Integrity serves as a filter on associativism, but only in the context of this internal standpoint. We can only exclude bad associations insofar as bad means that its actions fail ‘as action’. According to Korsgaard, action fails to live up to the standards of what it is to be an action if it fails to be authorised and owned. And for the ownership of action we require that it could be accounted for in terms of a coherent reason or intention.⁷¹ In this sense, integrity screens bad associations insofar as these associations are ‘badly constructed’ rather than morally flawed.

There is no reason for such modesty, if, as Korsgaard believes, the internal standards for what is and is not an action would *coincide* with the external standards for judging the morality or wrongness of an action. The upshot of this view is precisely that evil stems from the internal defectiveness of action ‘as action’. Morality is grounded in the free will, wickedness in subordination of the free will. To put it differently, willing and acting in accordance with the moral law are simply what it means to be an agent capable of willing and acting at all. We have to be moral, for this is what it takes to be fully autonomous. Now, if this view were true, then we could claim to have completely solved the problem of bad associations.

However, I have already argued that this view of things does not hold in the case of individual persons.⁷² The same objections apply to this bolder version of associativism. While it is true that integrity demands that the state act on laws or maxims that are authoritative for its own actions, there is no reason to believe that integrity should further require that the content of these laws and maxims should be such that they could reasonably be willed to guide the choices of states in general. Moreover, if the immoral nature of an action would spring from its defect ‘as an action’, then the immoral actions of a state could never be owned by that state. Conceiving of integrity in this way may lead to a moral vacuum in which it is impossible to hold any agent responsible for immoral deeds. We should, therefore, keep these internal and external standpoints apart.

⁷¹ See Chapter II, Section 4 above.

⁷² In Chapter II, Section 4.

Integrity is a procedural virtue emphasising that associations should be constituted on the basis of the coherence and unity provided by rule-governed order. Integrity is about the rule of law or principle. The *content* of these laws and principles, that is, their moral rightness or reasonableness, is a different matter altogether.

To sum up, value-dependent associativism of the sort that I am defending here provides at least some filter whereas identity-dependent associativism either resorts to exogenous provisos or offers nothing at all. That said, we have to recognise that unjust associations may still create *prima facie* associative obligations. And this may still sound absurd. In order to explain why this is not entirely absurd, the value-dependent approach needs to show that even the integrity of unjust associations, like the Mafia, is a virtue.

Generally, integrity is an independent virtue partly also because it allows us to understand how actions are attributable to persons who have authorised them. Integrity explains, among other things, how it is possible that there are continuous agents who can be held to account for what they have done in the past. Now, a person's character may be substantively flawed in the sense that this person adheres to maxims that ought to be rejected from the moral point of view. But this does not entail that the integrity of that person is no longer a virtue or a value. Take the example of Donny, a Mafioso, who assassinates Sonny. Clearly, this constitutes a moral wrong. Moreover, Donny appears to be a person of highly dubious morality as his practical identity is organised according to the principle of selfishness and the laws of the mob. Now, while he ought to renounce his mobster principles, it is still arguable that the bare fact that Donny has integrity counts for something. Despite of what he has done, indeed, precisely because of what he has done, most of us think it important that we are able to treat Donny as an agent. After all, we think it proper that he be held accountable for his earlier actions, receive punishment, show a sense of remorse, and so forth. So, all other things equal, it is in a way still more preferable that Donny commits this crime, not a loose collection of

intentional episodes. If, on top of this crime, Donny had no integrity, if he were temporally discontinuous in the way that the extremist imagines, then accountability would inevitably slip through the cracks of time.

In other words, integrity is a virtue distinct from justice, because it allows us to assign responsibility where responsibility should be assigned. It facilitates a mode of agency. This is essentially no different in the case of associations. If a state, say, wages war on its neighbours under the principle of collective selfishness, we may rightly describe this state as unjust. But this does not mean that the associative relatedness or the integrity constitutive of this state is entirely without value. It is precisely because a state possesses a degree of integrity that we can treat it as a unified and continuous agent. In contrast to an unruly multitude, we can assign responsibility to corporate persons, and insist later, for example, that sanctions be levied or reparations be paid. All things equal, it is still better to face an unjust state than a loose band of robbers. The value of integrity is not concurrent with the importance of justice. Justice is one thing, accountability and moral personality another. If, then, associative obligations are grounded in the value of integrity, it is not absurd to suggest that there could be prima facie obligations even where they clash with morality. These obligations may consequently have to be overruled by other considerations, but integrity is a virtue whatever a person does.

Finally, this leaves the objection that we have to make do with a form of value pluralism. Associativism argues that associations and relations have a basic value such that membership could directly create associative obligations. At the same time it may have to be acknowledged that there are duties that we owe to all human beings. Put differently, associative obligation could come into conflict with the wider concerns of justice or morality. Moreover, our obligations to one association may collide with our obligations to others. What we have to come to terms with, then, is that political obligation operates on the basis of an all things equal condition. And whether in the

final analysis a particular obligation should be decisive or trumped by other obligations or concerns, may be difficult to say. For example, if I discover that my good friend Donny is planning to kill Sonny, intuition suggests that the duty of justice should outweigh my duties of loyalty as a friend. But what do I say upon learning that Donny is involved in embezzlement and tax fraud? Do I gather evidence by covertly recording our conversations, and then make an anonymous phone call to the Inland Revenue? Such hard cases present us with profound dilemmas.

Now, from an ideal point of view it might have been expected that political theory could harbour no room for a plurality of values and obligations. Ideally, we have a full set of priority-rules, and ultimately a theory of everything, capable of telling us what we ought to do even in these cases. But I doubt that this ideal is much more than a pipedream. It should be recognised, in my view, that in many of such cases the dilemma is not apparent but profoundly real and persistent. We can take the formal impartialist's advice and ask ourselves whether or not it is 'reasonable' for anyone to reject this or that principle, but the fact is that at some point we simply run out of intuitions to work with. At times we simply do not know what 'reasonable' means, and which obligation should overrule the other. In my view it is far more preferable to appreciate that practical reason harbours this dilemma than to conjecture theoretical harmony at the expense of getting stuck with the answer that there are no political obligations at all.

Why is the dilemma deep and persistent? To conclude, I offer a brief answer. Throughout this thesis, I have been concerned to show that there is a kind of reasoning and obligation that is constitutive of a particular mode of agency. This is not just the case for corporate persons – we could, in fact, leave the idea of corporate personality aside here – but also for the personal identity of individuals. As we saw earlier, the notion of personal identity refers to a kind of connectedness or association. To be a temporally extended person or agent is merely to recognise a kind of associative obligation to your past and future selves. I further argued that these obligations are

primitive because the mode of agency that they facilitate happens to be an essential part of our way of being, our intuitions and judgements. Now, apart from this kind of associative obligation, it may be true that morality places us under a different sort of duty, to wit, a duty that we owe to other persons. However, while I do not want to affirm or reject such moral duties, it is I think obvious that ethical individualism has no choice but to affirm the associative obligations that are constitutive of the personal identity of individuals. At least insofar as theories are based on the notion of the individual as the bearer of interests and responsibilities, they *presuppose* these associative obligations. Thus, it is not surprising to see dilemmas emerging, even without having introduced wider obligations to the state or other associations. Since individualists claim that we are bearers of interests and responsibilities, they clearly rely on the idea of personal identity. And, as we have seen, to rely on the idea of personal identity, is to corroborate associative obligations and other normative considerations that derive, not from the higher ground of impartial duty, but directly from the relational fact of identity itself.

Conclusion

In this chapter, I compared two basic theories of associative obligation. *Assimilationism* argues that principles, thought to be universally 'non-rejectable', may nonetheless sanction political obligations with a narrower scope. Voluntarism seems a prime candidate for this job, but in the non-contractual setting of states this merely implies that there are no associative obligations. A better prospect is offered by combining the imperative of justice with an acute sense for administrative efficiency. I have argued, however, that for several reasons this attempt is problematic and likely to collapse.

Associativism presents a natural alternative by arguing that associations have obligation-generating significance. Associativism underpins this argument by appealing

to the claim that political obligations are conceptually inherent in the concepts of law, citizenship and state. However, while this claim is correct, we need an additional argument to account for the normativity of these obligations. Many associativists appeal to identity-based arguments. However, in contrast to these approaches, I argue that value-dependent arguments are more persuasive. In addition, while these value-dependent arguments tend to focus on nationality and citizenship, I have suggested an account emphasising the virtue of personal integrity.

In sum, I wanted to demonstrate that there are reasons to avow corporate personality as an irreducibly normative idea. Chapter V put forward the hypothesis that the corporate model is most plausible when placed in the context of our connectedness with past and future generations. I claimed it to be widely acknowledged that future generations matter, that individualist theories of intergenerational justice fail to account for this significance, and that the idea of corporate personality, when cashed out in terms of transgenerational association, addresses this issue more successfully. The same line of argument, this chapter suggests, can be directed towards the past. I argue that political obligations and responsibilities are handed down by inheritance, that these inherited obligations are not adequately captured by ethical individualism, and that corporate personality provides us with a more promising model.

CONCLUSION AND SUMMARY

This thesis argues that human associations qualify as moral persons in and of themselves. Most political theorists are firmly committed to the principle of ethical individualism. On this view, associations have no independent moral standing. Their ethical importance derives solely from the standing of their individual members. This thesis maintains, in contrast, that we need to revive corporation theory. We should return to this theory, because the notion of corporate personality, unlike the communitarian appeal to culture and language, is able to explain why associations matter in a fundamental way. Like ethical individualists, I believe that persons matter fundamentally. But there are no good reasons to hold that only individuals are persons.

Ethical individualism seems to derive mainly from the ontological prejudice that groups are reducible to their individual members in a way that individuals are not. Groups do not have a will or interest, at least not separately from their individual members. And so there appears to be little room for corporate personality. Essentially, this thesis shows that the case for ethical individualism is never that straightforward. The philosophical catch is that the individual, when viewed in a diachronic context, is just another kind of association. Individual persons are merely bundles of psychological states, as Parfit reminds us. To be sure, this fact need not stop us from claiming that individuals have responsibilities and rights as unified agents, but the basic question shifts from: 'Could relational facts have obligation-generating significance?' to 'Why do *these* relations have ethical standing, not others?'

The gist of my argument is that corporate personality could be revived by constructing it as a normative theory. This reconstruction directs us to the ideas of associativism and integrity, which are in fact interchangeable with the notion of corporate personality. It is due to their integrity that individuals and associations become single agents, capable of willing and acting as a unified whole. Integrity is a virtue both for individuals as well as for groups. It captures an important standard of

associative conduct, which I claim is constitutive of our agency, while distinct from justice or morality. Associativism is the view that relational facts have a basic value and directly impose obligations on their members. It finds support in the idea that corporate personality elongates our ethical horizons. Corporate personality, I argue, has the unique ability to connect past, present and future generations in meaningful and practically significant ways.

Let me spell out the entire argument of this thesis in greater detail. Essentially, corporation theory contends that associations should be regarded as the independent bearers of interests, rights and obligations. That is to say, the interests, rights and obligations of the association are directly reason-giving, without needing a further justification in terms of the interest or good or will of the individual. The notion of a corporate person further entails that associations are capable of unified agency.

This notion of corporate personality is traditionally thought to depend on one of the following two justifications. According to the *fiction theory*, a group is a corporate person if certain people or institutions are prepared to feign its personality. Ultimately, associations do not have interests and intentions that exist separately from individual interests and intentions. We merely pretend that they do. According to *corporate realism*, the personality of a group depends, not on the acceptance of a fiction, but on its own reality as an agent. In other words, associations have real interests and intentions. Their personality is not just made up, but comparable to that of individuals. Like the latter, the group qualifies as a right-and-duty-bearing unit on the grounds that it forms an organism with a life and a mind of its own.

The problem is that each of these justifications faces serious objections and limitations. The fiction theory, on the one hand, is of no use to ethical and political theory. On the basis of fiction we could personify anything. Corporate realism, on the other, is simply not plausible. At any rate, the latter view cannot give corporation theory a new lease of life.

Given these limitations, it is hardly surprising that the notion of corporate personality has almost disappeared from the debate. In the main, contemporary political theory is based on the principle of ethical individualism. The individual is the basic unit of analysis in theories of rationality and morality. All conceivable reasons for action – the standards of both the right and the good – are claimed to derive from an action's impact on individual well-being. What is more, only the individual is thought to be genuinely capable of assuming responsibility for his actions. Ethical individualism could still argue that associations or communities play an important role in our lives. But their value is at best instrumental to the individual good. Associations are never a good in and of themselves.

Is corporate personality dead? It is widely believed that this is indeed the case. However, this thesis shows there are still ways to breathe new life into this notion. In order to accomplish this, we need to break with the received dogmas that view group personality either as a fiction or as a separately existing collective mind. We need a different kind of justificatory theory. While this theory cannot appeal to fiction, the main challenge is to show that, at the same time, it manages to steer clear of corporate realism. Another way of putting it is to say that we have to reconcile the notion of a corporate person with a reductionist or individualist ontology. We have to argue that associations qualify as moral persons in the sense set out above, but acknowledge that groups do not have separate minds.

How can this paradox be unravelled? Part of the answer involves a switch from an *ontological* to an irreducibly *normative* conception of the person and personal identity. This normative conception has in fact been developed by ethical individualists themselves. But I go on to claim that it can be employed against individualism itself. However, before discussing that issue, it is important to understand that the paradox does not only emerge in the case of corporate personality. Although this is often overlooked, it also surfaces at the level of individual personality.

Chapter I explains why this is the case. The idea is that individuals and associations are analogous after all. However, in contrast to what corporate realists have always claimed, this analogy is argued to hold up *inversely*. Groups are not individuals 'writ large', but individuals are groups 'writ small'. This inversed comparison emerges from a reductionist theory according to which the identity of a person consists in the psychological continuity and/or connectedness between its temporal stages. On this view, there is no indivisible Ego, which exists separately from this bundle of psychological states and is sometimes said to be its owner. Personal identity is purely relational. Whether or not I am identical with a future self is a matter of degree, of stronger or weaker psychological relatedness.

Now, it will be objected that by inverting the analogy between individuals and groups we have only made a bad situation even worse. For the bundle theory seems to imply that we have to abandon the notion of the person altogether. Indeed, according to the 'extreme claim' we should not just reject the appeal to the interests and good of associations, we should also reject the appeal to the temporally extended interests and good of individual persons. Far from having rescued the idea of corporate personality, we have only damaged ethical individualism. What we are left with is a highly atomised brand of consequentialism, according to which the importance, not just of social groupings, but even of personal identity is at best derivative.

However, this is not the only interpretation of the bundle theory. In Chapter II, it is argued that ethical individualism could reject the extreme claim. For this to be successful, ethical individualism has to modify its understanding of the role of the person and personal identity in normative theory. Assuming that a return to the Self is not a plausible option, the key is to advance an understanding of personal identity, not just as a purely *psychological* bundling of mind-states, but also as a form of connectedness that is inherently *normative*. According to the moderate claim, for example, the relatedness between the stages of a person is partly constituted by a

concern for each other's well-being. Being connected to a future self *is* to feel a kind of prudential concern. And being connected to a past self in the manner of a person *is* to accept past commitments, intentions and actions as your own. In short, while the moderate claim acknowledges that personal identity is essentially a relation, this relational fact gives us reasons for action.

Why should we support the moderate rather than the extreme claim? Two arguments are important here. First, we have to concentrate on the concept of a person in general and its role in normative theory. One way of approaching this subject suggests that theories of the person and personal identity provide normative theories with an objective foundation. However, this approach fails to help us out. The problem is that both claims are at odds, not over ontology, but over the question whether psychological bundles have an inherent and basic normative significance. The second way of approaching this issue – I call it the *practice theory* – suggests that theories of the person and personal identity are the subject of normative theorising. On this view, normative theory operates on the basis of reflective equilibrium. If, then, we apply the practice theory to the dispute between the extreme and moderate claims, the former seems very implausible. The idea that every stage of a person forms a separate agent as such, with interests, duties and rights of its own, simply does not correspond to our way of life. We all recognise that we are in a practically relevant sense deeply connected with our past and future lives.

Second, we have to give more depth to the idea that personal identity is partly defined by normative connectedness. I therefore look at the notion of *integrity*. It is argued – first only from a synchronic point of view – that in order to be a unified agent at all, a person needs to be able to organise his beliefs, desires and choices on the basis of reasons. That is to say, as a unified agent, my actions must be explicable in terms of general principles or policies that are authoritative over all my actions. I must be able to justify or account for my choices and actions in terms of their internal coherence. Three

further points should be noticed. First, integrity does not necessarily coincide with morality. It is not the case that in order to be a *unified* agent one must necessarily be a *moral* agent. Second, unified agency or personality turns out to be an end, a goal or an achievement. It is something we should do our best to attain, that we could attain to certain degrees. Third, the notion of integrity can be seen to apply at a diachronic level. It offers a particular conception of personal identity over time.

With the practice theory and this conception of integrity in place, the argument then returns to corporate personality. I claim its possibility is demonstrated by extending the moderate claim to the level of nations or states or other groups. This extension of the moderate claim could be referred to as *associativism*. Roughly, associativism suggests that *interpersonal* forms of connectedness could have the same normative significance as the *intrapersonal* connectedness between past, present and future selves. While associations are quite clearly relational facts, this theory goes on to argue that they still directly impose obligations on their members. For example, it argues that the communal good, like the good of a temporally extended person, still has non-derivative weight in our practical reasoning.

But why should we accept associativism rather than ethical individualism? Fundamentally, our task is to use the practice theory and the related notion of integrity to build a case for associativism, a case that therefore mirrors the justification for the moderate claim. Building on the practice theory, we first have to show that associativism and the idea of corporate personality are normatively attractive when tested against our well-considered judgements and intuitions. Moreover, corporate personality has to be capable of doing justificatory work that ethical individualism leaves undone. Second, we have to show that the notion of integrity could also be applied at the wider level of associations like nations and states.

Chapter III begins with the last part of that task. Corporate personality, most people agree, must be consistent with ontological reductionism or individualism about

collectives. This means that we have to understand the notion, not as a mysterious Group Ego, but as a bundling of the intentions and preferences of individuals. To put it differently, a group-person emerges when a number of individuals come to *share a collective* intention or preference, just as the personality of an individual emerges when a number of person-stages come to *share a future-directed* intention or preference. Above all, what this requires is that the associates think in a particular mode of thought, containing two levels. On the first level, so-called collective intentions (or purposes or preferences) must be defined over the actions of a number of individuals (and their outcomes). Collective intentions relate to things or ends that we could do or achieve *together*. Second, instead of basing their decisions purely on individual intentions, those individuals who think in this mode of thought take these collective intentions to be authoritative over their choices. When a number of people thus come to hold collective intentions that are sufficiently overlapping, an association may be correctly said to have personal integrity, just as an individual may be said to have personal integrity when a number of person-stages come to hold individual intentions that are sufficiently overlapping.

Since this approach to corporate personality is clearly reductionist in an ontological sense, care has to be taken to prevent the model from collapsing back into a form of ethical individualism. We could distinguish at least three ways in which such a collapse is likely to occur. *Conventionalism* argues that collective intentions and preferences could ultimately be broken down in individual intentions and preferences. *Contractualism* argues that associativism must be subjected to the normative test that no-one should be allowed to be worse off as the result of his participation. Finally, *voluntarism* argues that all members need to have freely given their consent to becoming an associative member. In opposition to these three interpretations, it is argued respectively that collective intentions are basic, that whether or not intentions or

preferences are shared in the above sense is an empirical matter, and that while voluntary consent may be a sufficient condition, it is not a necessary condition.

This idea of corporate personality is thus designed to remain within two side-constraints. First, it does not lapse into an implausible kind of anthropomorphic metaphysics, which earlier theories relied on. So it fits the *reductionist* constraint. Second, it is nonetheless able to make an independent contribution to normative theory. That is to say, it fits what I call the *sovereignty* constraint. I next take up a different challenge. As it stands, the model of associative integrity and shared intention developed in Chapter III seems applicable mainly to small-scale, deliberative and egalitarian contexts. Chapter IV aims to modify this account of corporate personality such that it could more plausibly be applied to larger human associations and to the state in particular. In other words, the question is whether or not this notion of a corporate person could also be made to fit a third constraint, which is *empirical* in nature.

Essentially, the problem with the idea of the personality of the state is that it projects a collective intention that is shared at a societal scale. Such a projection is problematic for at least three reasons. First, states usually include a vast number of people and many intricate and elaborate forms of collective activity, co-ordination and planning. So there is a clear problem with complexity. That is, we simply cannot expect all citizens to be consciously aware of and guide their actions according to a collective or executive intention at this level of complexity. Second, states are often sharply demarcated by territorial boundaries. If states are corporate persons, a collective intention would have to be exclusively shared by a population or membership that can be defined territorially, a requirement that seem empirically vulnerable. Third, states claim a kind of authority that is alien to most other associations. While associations can in principle co-exist with other, like-minded associations, states are monopolies. The problem, then, with the notion of the integrity of the state is that it would seem to rule out pluralism, dissent and difference. Yet, as has often been argued, it is a fact that most states harbour a great

many different practical and cultural identities. Given this 'fact of pluralism', it is difficult to see how states could have integrity in the sense defined above.

In response to these challenges, I attempt to revise the model so as to accommodate the complexity, territoriality and internal pluralism of states. As a first step, I argue that we should think of intention as an *interpretive* rather than a *psychological* concept. According to the latter concept, a collective intention should always be thought of as a conscious state in the minds of all those who are members. Each citizen must assume the executive point of view on what they as a citizen-body aim to achieve together. According to the interpretive approach, however, this need not be the case. What is important, on this view, is merely that our actions are sufficiently 'meshing' such that a collective intention could later be constructed. Collective intentions, it suggests, are *implicit* in what we do. They manifest themselves only by the 'constructive interpretation' of a coherent practice. The principle idea is that we start from a historically evolved pattern of collective action, and then try to account for the practice as a whole in terms of one single purpose or principle. To put it differently, while none of the participants actually needs to assume the executive point of view, the state has integrity when the actions of its citizen-body are justifiable on the basis of a single principle or rationale.

However, while this interpretive approach seems to assuage the problem of complexity, it does little to alleviate the problems of boundaries and pluralism. It is important, therefore, to also introduce an institutional dimension to the model. Essentially, the idea is here that the corporate personality of the state could be facilitated by intermediary rules of recognition or constitutive rules. Such rules shift the *locus of integrity* to a different level. Notice that even though the idea of integrity is interpretive, that is, immanent in a set of overlapping or meshing activities, it should still be effective as a normative constraint on that practice. The notion of a locus goes on to specify where precisely integrity should be effective as such a constraint. On a non-institutional

approach, a collective intention must be normative for the population or citizen-body at large. However, once the appropriate rules of recognition are in place, the locus of integrity shifts from the population at large to the official bodies of the state, which legislate and adjudicate on behalf of the community. By equipping corporate personality with this institutional dimension, we are thus in a much better position to deal with the problems of boundaries and pluralism.

The remainder of the thesis turns to normative issues. Like personal identity, corporate personality is not so much an ontological theory as a theory of practical reason. It must therefore demonstrably have practical merits. We could call this the *practicality* constraint. Chapter V investigates a number of hypotheses to this effect. First, it briefly discusses and calls into question the communitarian thesis that culture and language are irreducibly social or communal goods. Second, I look at the idea that associativism extends our concern from a narrow focus on the present, to our temporally extended good, and then to an even wider concern for the good of others. On this view, corporate personality suggests that we have grounds to treat the good of our fellow-associates as reasons, just as the idea of an individual person suggests that you have grounds to treat your future interests as basic. To put it differently, associativism provides the grounding for norms of justice and civility without which the social order could not endure. However, this hypothesis remains debatable, mainly for the reason that ethical individualism supports powerful rival explanations. On the one hand, there is the Hobbesian view that norms of reciprocity or justice or civility are mutually advantageous to each of the individuals involved. On the other hand, there is the Kantian view of justice as impartiality, that is, the idea that each individual agent is entitled to an equal and fundamental respect as a matter of pure morality.

To avoid these objections, I introduce a more specific variant of this general hypothesis. Corporate personality, I argue, extends our ethical horizons, not only to others, but also to past and future generations. And here, I argue, neither justice as

mutual advantage nor justice as impartiality provides a persuasive alternative. To demonstrate the shortcoming of justice-based accounts, I begin by focussing on the interests of future generations. Theories of justice often claim to be able to show that we are under a duty of justice to take these interests in consideration. However, as has been acknowledged before, these theories fail to make good on that pledge for a variety of reasons.

Chapter VI furnishes the positive half of the argument. I argue that corporate personality extends across history in such a way that generations of associative members could be compared to the stages of a person. Ethical individualism, on the one hand, conceives of each generation as a sovereign nation, in control of its own destiny and unburdened by the sins of its ancestors. Since generations are viewed as separate nations, the reason why present generations care about the good of future generations could only be derived from a conception of justice. Corporate personality, on the other hand, holds that all generations belong to the same association. In fact, this conception of a transgenerational association allows us to conceive of the issue in terms of the *goodness* of a life taken as a whole instead of *justice* between separate lives. To further elaborate on this idea, I compare a maximisation criterion with certain narrative or diachronic criteria for the intrapersonal distribution of well-being. I suggest that the latter, especially, could be fruitfully applied to transgenerational associations.

Finally, Chapter VII seeks to expand this temporal argument for corporate personality by looking at the past. It has often been argued that political obligations are not freely assumed, but inherited. Similarly, present generations are bound by constitutions, laws and treaties that are often well over a century old. Still, liberal individualists would agree that we are normally bound by such obligations and commitments. At the same time, the validity of such obligations seems clearly contestable. How could political obligations be passed on down the generations without our consent? Again, ethical individualism is committed to the idea of generational

sovereignty. On this view, the only available explanation of political obligations and commitments could be found in an appeal to justice. Natural justice theories argue that political and associative obligations derive from our duties of universal morality. The general idea is that we need political associations and institutions in order to be able to effectively realise justice. And insofar as the institutions of a particular state are indeed just, the associative obligations of its citizens are best understood as a specific manifestation of their moral duty to humanity at large. But this derivation from justice, too, has its problems. One well-known problem is that many of our associative duties seem to call out for an independent basis in a way that assimilation with duties of justice rules out. Another problem is that in many cases natural justice theories fail to deliver the goods. At any rate, it is argued that a more plausible explanation could be found in the idea of corporate personality. On this view, generations are part of one transgenerational association. Again, this conception allows us to compare the relatedness between generations with the relatedness between the stages of one person.

I should, as a final point, say something about the limitations of this thesis. First, it does not address the empirical issues that could be raised by the idea of corporate personality. Most conspicuously, perhaps, I provide no general classification of all the kinds of association that might qualify as corporate persons in contemporary societies. Chapter IV focuses on the personality of the state, but I say little about associations like the family, voluntary associations, trade-unions, cultural groups or more loosely organised groupings such as social-economic class, mobs and crowds.¹ Nor am I interested in questions as to whether this or that particular association qualifies as a continuous person or not. These questions will no doubt arise at some point, as sometimes they will also arise for individuals, but they are not decided here. I confine myself to using broad strokes, not because these things are unimportant, but because

¹ For the position of mobs and crowds, see L. May, *The Morality of Groups: Collective Responsibility, Group-based Harm, and Corporate Rights* (Notre Dame: University of Notre Dame, 1987). The origins and the plurality of human group life are investigated in A. Black, *State, Community and Human Desire*.

corporation theory requires a new philosophical justification. Before we can confront these empirical issues we need to rediscover new conceptual space. And to find this space, we have to re-examine the idea of a person and its position in ethical and political thought.

Second, this thesis is about corporate personality; it does not explore the issues raised by the related theory of associative pluralism, as it flourished in Britain at the start of the Twentieth Century. According to Nicholls, this school, with among its leading figures people such as Figgis, Cole and Laski, advanced three basic arguments, each pertaining to the appropriate relation between the state and other associations within its territory.² First, it argued that liberty was the greatest end in politics. Second, it advanced the notion of the real and moral personality of associations. Third, it maintained that the state had no greater authority over its members than other associations. Thus, pluralism was highly critical of the state and the notion of sovereignty, as it figured in the theories of Hobbes and Rousseau for example. I should point out that I am interested only in the second claim. I argue in Chapter IV that the state, although clearly different in many ways from other associations, has its own personality and unity. Furthermore, although I do not press this claim here, it seems to me that some notion of sovereignty cannot be avoided. Of course, how precisely modern, representative democracies should be organised internally to give sufficient liberty and weight to associations that reside within their higher order unity remains a pertinent question.³ But it is not the topic that this thesis addresses. I am interested mainly in the question how associations in general should be ethically appraised, in particular in view of the challenge presented by ethical individualism.

Third, the idea of the corporate person of the state has been claimed to raise the spectre of a kind of collectivism that seems unacceptably hostile to individual rights and

² D. Nicholls, *The Pluralist State: The Political Ideas of J. N. Figgis and his Contemporaries*, second edition (Basingstoke: St. Martin's Press, 1994).

³ For a recent attempt to revive pluralist thinking, see P. Hirst, *Associative Democracy: New Forms of Economic and Social Governance* (Cambridge: Polity Press, 1994).

liberties. For example, apart from making a valid theoretical point, Popper was also, perhaps even mainly, concerned with bringing the theories of Plato and Hegel and others in political disrepute. Like other liberals of that period, he linked corporation theory, often in a loose and entirely unconvincing manner, to fascism, communism, racism and the glorification of war.⁴ However, it is important to see that the corporate notion of a person – at least in the shape that I have given to it here – does not make light of the ethical significance of the individual. Rather, it suggests that both notions co-exist. We still have prudential and moral reasons for action that are entirely justifiable in terms of our own interests and preferences and those of other individual agents. But my point is that there is also a category of reasons that is ultimately only explicable on the basis of collective interests, preferences and associative membership. It seems to me a fundamental mistake to think that we are ultimately *either* corporate *or* individual persons. We could have reasons qua and duties to individuals *and* reasons qua and duties to groups. To put it differently, what emerges from the discussion below is the view that the unit of a person is not fixed either at the collective or individual level, but more flexible and adaptable. A strict dichotomy between individualist and collectivist theories should in any case be dismissed. Co-existence need not be peaceful co-existence. And the pluralism that results on this view may well lead to deep practical conflict – for example, between our moral and associative duties – and so to the further question of its resolution. But that question, too, goes beyond the scope of my argument. What I want to demonstrate is that this conflict could exist. Not all our reasons are reducible to the goods and preferences of individual persons. The point, then, is not that we ought to reject individual rights and liberties, but that ethical individualism has limitations. And it is these limitations that lead us back to the notion of corporate personality, and towards a greater sense of balance between the two.

⁴ K. R. Popper, *The Open Society and its Enemies*.

Associations, in short, should qualify as persons because they accommodate a kind of temporal continuity that individual persons cannot provide. Associations do not only widen our ethical horizons, they also *lengthen* them to embrace the past and the future. If personal identity supplies the ground for longstanding projects or relationships between individuals based on love, friendship and reciprocity, the wholeness of corporate persons furnishes even longer standing projects and relationships. That is to say, we need corporation theory to obtain a handle on those projects and relationships that stretch out beyond a single generation. Whereas individual rights, obligations and responsibilities are destined to remain generationally fractured, corporate personality holds out the possibility of a genuinely transgenerational ethic, and with that, the accumulation and conservation of the knowledge, capital and resources on which human civilisation is built. Associations do not have to start from a clean slate once one generation of members has been replaced by the next; they accumulate achievements and pass them down the generations. In other words, corporation theory furnishes the grounds for the fulfilment of exactly those human goods, those plans and ambitions that require more than a lifetime's effort. I shall not waste time by stating the self-obvious, that many great human aspirations are of this nature, how much we are ourselves the produce of historical evolution. It is perhaps less obvious, though, that only in the persisting nature of associative ties we find the possibility of evolution and history in the first place.

BIBLIOGRAPHY

- Adams, R. M., 'Should Ethics Be More Impersonal?', in: Dancy, J., (ed.), *Reading Parfit* (Oxford: Blackwell, 1997), pp. 251-289.
- Ainslie, G., *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person* (Cambridge: Cambridge University Press, 1992).
- Aristotle, *The Politics*, in: Everson, S., (ed.), *The Politics and the Constitution of Athens* (Cambridge: Cambridge University Press, 1996).
- Axelrod, R., *The Evolution of Social Co-operation* (New York: Basic Books, 1984).
- Bacharach, M., 'Interactive Team Reasoning: A Contribution to the Theory of Co-operation', *Research in Economics*, Volume 53, No. 2 (June 1999), pp. 117-147;
- Ball, T., 'The Incoherence of Intergenerational Justice', *Inquiry*, 28:3 (1985), pp. 321-337.
- Barker, E., 'Introduction', in: O. Gierke, *Natural Law and the Theory of Society*, E. Baker (ed.), (Cambridge: Cambridge University Press, 1934).
- Baron, M., 'Impartiality and Friendship', *Ethics* 101 (1991), pp. 836-857.
- Barry, B., 'Circumstances of Justice and Future Generations', in: Sikora, R. I., and Barry, B., (eds.), *Obligations to Future Generations* (Cambridge: The White Horse Press, 1996), pp. 204-248.
- Barry, B., 'Intergenerational Justice in Energy Policy', in: Maclean, D., and Brown, P. G., (eds.), *Energy and The Future* (Totowa: Rowman and Littlefield, 1983), pp. 15-30.
- Barry, B., 'Self-Government Revisited', in: *Democracy and Power: Essays in Political Theory 1* (Oxford: Clarendon Press, 1991), pp. 156-186.
- Barry, B., 'Sustainability and Intergenerational Justice', in: Dobson, A., *Fairness and Futurity* (Oxford: Oxford University Press, 1999), pp. 93-117.
- Barry, B., *Justice and Impartiality* (Oxford: Clarendon Press, 1995).
- Barry, B., 'Justice Between Generations', in: *Liberty and Justice: Essays in Political Theory 2* (Oxford: Clarendon Press, 1991), pp. 242-258.
- Barry, B., *Theories of Justice* (London: Harvester Wheatsheaf, 1989).
- Black, A., *State, Community and Human Desire* (Hemel Hempstead: Harvester Wheatsheaf, 1988).
- Blackburn, S., 'Has Kant Refuted Parfit?', in: Dancy, J., (ed.), *Reading Parfit* (Oxford: Blackwell Publishers, 1997), pp. 180-201.
- Bosanquet, B., *The Philosophical Theory of the State*, fourth edition (London: MacMillan, 1958).
- Bradley, F. H., 'My Station and Its Duties', in: *Ethical Studies*, second edition (Oxford, Clarendon Press, 1927).
- Bratman, M. E., 'I Intend That We *J*', in: *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), pp. 142-161.
- Bratman, M. E., 'Identification, Decision, and Treating as a Reason', in: *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), pp. 185-206.
- Bratman, M. E., 'Intention and Personal Policies', *Philosophical Perspectives*, 3 (1989), pp. 443-469.
- Bratman, M. E., 'Planning and Temptation', *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), pp. 35-57.

- Bratman, M. E., 'Reflection, Planning, and Temporally Extended Agency', *The Philosophical Review*, Vol. 109, No. 1 (January 2000), pp. 35-61.
- Bratman, M. E., 'Shared Cooperative Action', in: *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), pp. 93-108.
- Bratman, M. E., 'Shared Intention and Mutual Obligation', in: *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), pp. 130-141.
- Bratman, M. E., 'Shared Intention', in: *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), pp. 109-129.
- Bratman, M. E., 'Toxin, Temptation, and the Stability of Intention', in: *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), pp. 58-90.
- Bratman, M. E., *Intention, Plans, and Practical Reason* (Cambridge: Harvard University Press, 1987).
- Bratman, M. E., 'Hierarchy, Circularity, and Double Reduction', in: Buss, S., and Overton, L., (eds.) *Contours of Agency: Essays on Themes from Harry Frankfurt* (Cambridge MA: The MIT Press, 2002), pp. 65-85.
- Braybrooke, D., 'The Social Contract and Property Rights across the Generations', in: Laslett P., and Fishkin, J. S., (eds.), *Justice Between Age Groups and Generations* (New Haven: Yale University Press, 1992), pp. 107-126.
- Brink, D. O., 'Rational Egoism and the Separateness of Persons', in: Dancy, J., (ed.) *Reading Parfit* (Oxford: Blackwell Publishers, 1997), pp. 96-134.
- Brink, D. O., 'Self-love and Altruism', *Social Philosophy and Policy*, Volume 14, Number 1 (1997), pp. 122-157.
- Brecher, B., 'Our Obligation to the Dead', *Journal of Applied Philosophy*, Volume 19, No. 2 (2002), pp. 109-119.
- Broad, C. D., 'Self and Others', in: Cheney, D. R., (ed.), *Broad's Critical Essays in Moral Philosophy* (London: Allen & Unwin, 1971), pp. 262-282.
- Buchanan, A. E., 'Assessing the Communitarian Critique of Liberalism', *Ethics* 99/1989, pp. 852-882.
- Buchanan, A. E., 'Liberalism and Group Rights' in: Coleman, J. L., and Buchanan, A. E., (eds.), *In Harm's Way: Essays in Honor of Joel Feinberg* (Cambridge: Cambridge University Press, 1994), pp. 1-15.
- Buchanan, J. M., *The Limits of Liberty: Between Anarchy and Leviathan* (Chicago: University of Chicago Press, 1975).
- Burke, E., *The Reflections on the Revolution in France*, J. G. A. Pocock (ed.) (Indianapolis: Hackett Publishing Company, 1987).
- Butler, J., 'Of Personal Identity', in: J. Perry (ed.), *Personal Identity* (Berkeley: University of California Press, 1975), pp. 99-105.
- Caney, S., 'Individuals, Nations and Obligations', in: Caney, S., George, D., Jones, P., (eds.), *National Rights, International Obligations*, (Oxford: Westview Press, 1996), pp. 119-138.
- Carr, D., *Time, Narrative, and History* (Bloomington: Indiana University Press, 1986).
- Carr, D., 'Getting the Story Straight', in: G. Roberts (ed.), *The History and Narrative Reader* (London: Routledge, 2001), pp. 197-208.
- Carr, D., 'Narrative and the Real World: An Argument for Continuity', in: Hinchman, L. P., and Hinchman, S. K., (eds.), *Memory, Identity, Community: The Idea of*

- Narrative in the Human Sciences* (Albany: State University of New York Press, 1997), pp. 7-25.
- Carroll, N., 'Interpretation, History and Narrative', in: G. Roberts (ed.), *The History and Narrative Reader*, (London: Routledge, 2001), pp. 246-265.
- Chappell, T., 'Reductionism about Persons; And What Matters', *Proceedings of the Aristotelian Society*, Vol. XCVII (1997-1998), pp. 41-57.
- Cohen, G. A., 'Reason, Humanity, and the Moral Law', in: C. M Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), pp. 167-188.
- Cohen, P. R., and Levesque, H. J., 'Teamwork', *Noûs*, Volume 25, Issue 4 (September 1991), pp. 487-512.
- Constitution of the United States of America, 1787*, in: C. Rossiter (ed.), *The Federalist Papers* (New York: Mentor Book, 1999), pp. 509-530.
- Copp, D., 'Hobbes on Artificial Persons and Collective Action', *The Philosophical Review*, Volume 89, Issue 4 (October, 1980), pp. 579-606.
- Dagger, R., 'Membership, Fair Play, and Political Obligation', *Political Studies*, 2000, Vol. 48, pp. 104-117.
- Dahrendorf, R., *Reflections on the Revolution in Europe* (New York: Time Books 1990).
- Dan-Cohen, M., *Rights, Persons, and Organizations: A Legal Theory for Bureaucratic Society* (Berkeley: University of California Press, 1986).
- De-George, R. T., 'The Environment, Rights, and Future Generations', in: Goodpaster, K. E., and Sayre, K. M., (eds.) *Ethics and Problems of the 21st Century* (Notre Dame: University of Notre Dame Press, 1979), pp. 93-105.
- De-Shalit, A., *Why Posterity Matters: Environmental Policies and Future Generations* (London: Routledge, 1995).
- De-Shalit, A., *The Environment: Between Theory and Practice* (Oxford: Oxford University Press, 2000).
- Dennett, D., 'Conditions of Personhood', in: Rorty, A., (ed.), *The Identity of Persons* (Berkeley: University of California Press, 1976), pp. 175-196.
- Descartes, R., *Meditations on First Philosophy*, J. Cottingham (ed.), (Cambridge: Cambridge University Press, 1996).
- Dewey, J., 'The Historic Background of Corporate Legal Personality', *Yale Law Journal*, Volume XXXV (April, 1926), pp. 655-673.
- Dobson, A., *Justice and the Environment* (Oxford: Oxford University Press, 1998).
- Duff, P. W., *Personality in Private Roman Law* (Cambridge: Cambridge University Press, 1938).
- Dworkin, R., *Law's Empire* (Oxford: Hart Publishing, 1998).
- Dworkin, R., *Taking Rights Seriously*, seventh impression (London: Duckworth, 1994).
- Dworkin, R., 'Liberal Community', in: Avineri, S., and De-Shalit, A., (eds.) *Communitarianism and Liberalism* (Oxford: Oxford University Press, 1992), pp. 205-223.
- Elliot, R., 'The Rights of Future People', *Journal of Applied Philosophy*, Volume 6, No. 2 (1989), pp. 159-169.
- Elster, J., 'Interpersonal Choice and Political Thought', in: G. Loewenstein and J. Elster (eds.) *Choice over Time* (New York: Russell Sage Foundation, 1992), pp. 35-53.
- Elster, J., *Ulysses and the Sirens: Studies in Rationality and Irrationality* (Cambridge: Cambridge University Press, 1979).

- Elster, J., *Ulysses Unbound: Studies in Rationality, Precommitment, and Constraints* (Cambridge: Cambridge University Press, 2000).
- Feinberg, J., *Rights, Justice, and the Bounds of Liberty*, (Princeton: Princeton University Press, 1980).
- Fishkin, J. S., 'Limits of Intergenerational Justice', in: Laslett P., and Fishkin, J. S., (eds.), *Justice Between Age Groups and Generations* (New Haven: Yale University Press, 1992), pp. 62-83.
- Frankfurt, H. G., 'Freedom of Will and the Concept of a Person', *The Journal of Philosophy*, Volume 68 (January 14, 1971), pp. 5-20.
- Frankfurt, H. G., 'Identification and Externality', in: *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press, 1997), pp 58-68.
- Frankfurt, H. G., 'Identification and Wholeheartedness', in: *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press, 1997), pp. 159-176.
- French, P. A., *Collective and Corporate Responsibility* (New York: Columbia University press, 1984).
- French, P. A., *Responsibility Matters* (Kansas: University of Kansas Press, 1992).
- Gadamer, H. G., *Truth and Method* (London: Sheed and Ward, 1989), second revised edition.
- Garrett, B., *Personal Identity and Consciousness* (London: Routledge, 1998).
- Gauthier, D., *Morals by Agreement* (Oxford: Oxford University Press, 1986);
- Gauthier, D., 'Assure and Threaten', *Ethics*, Volume 104, Issue 4 (1994), pp. 690-721.
- Gauthier, D., 'Commitment and Choice: An Essay on the Rationality of Plans', in: Farina, F., Hahn, F., and Vannucci, S., (eds.) *Ethics, Rationality, and Economic Behaviour* (Oxford: Clarendon Press, 1996), pp. 217-243.
- Gauthier, D., 'Resolute Choice and Rational Deliberation: A Critique and a Defense', *Nous*, 31:1 (1997), pp. 1-25.
- Gewirth, A., 'Ethical Universalism and Particularism', *The Journal of Philosophy*, Vol. 85, No. 6 (June 1988), pp. 283-302.
- Gierke, O. von, 'The Nature of Human Associations', J. D. Lewis (ed.), in: *The Genossenschaft-theory of Otto Von Gierke* (Madison: University of Wisconsin Studies in the Social Sciences and History, 1935), pp. 139-157.
- Gierke, O. von, *Community in Historical Perspective*, A. Black (ed.), (Cambridge: Cambridge University Press, 1990).
- Gierke, O. von, *Natural Law and the Theory of Society*, E. Baker (ed.), (Cambridge: Cambridge University Press, 1934).
- Gierke, O. von, *Political Theories of the Middle Age*, F. W. Maitland (ed.) (Cambridge: Cambridge University Press, 1927).
- Gilbert, M., 'Agreements, Coercion, and Obligation', *Ethics*, Volume 103, Issue 4 (July 1993), pp. 679-706.
- Gilbert, M., 'Collective Remorse', in: *Sociality and Responsibility: New Essays in Plural Subject Theory* (Lanham: Rowman and Littlefield, 2000), pp. 123-140.
- Gilbert, M., 'Collective, Preferences, Obligations, and Rational Choice', *Economics and Philosophy*, 17 (2001), pp. 109-119,
- Gilbert, M., 'In Search of Sociality – Recent Developments in the Philosophy of Social Phenomena', in: *Sociality and Responsibility: New Essays in Plural Subject Theory* (Lanham: Rowman and Littlefield, 2000), pp. 154-164.

- Gilbert, M., 'Obligations of Joint Commitment', in: *Sociality and Responsibility: New Essays in Plural Subject Theory* (Lanham: Rowman and Littlefield, 2000), pp. 50-70.
- Gilbert, M., 'Reconsidering the 'Actual Contract' Theory of Political Obligation', in: *Sociality and Responsibility: New Essays in Plural Subject Theory* (Lanham: Rowman and Littlefield, 2000), pp. 97-122.
- Gilbert, M., 'The Idea of Collective Guilt', in: *Sociality and Responsibility: New Essays in Plural Subject Theory* (Lanham: Rowman and Littlefield, 2000), pp. 141-153.
- Gilbert, M., *On Social Facts* (London: Routledge, 1989).
- Godwin, W., *An Enquiry Concerning Political Justice*, in: Philp, M., (ed.), *Political and Philosophical Writings of William Godwin, Volume 3* (London: William Pickering, 1993).
- Goethe, J. W. von, *Egmont*, translated by Anna Swanwick, Vol. XIX, Part 3, The Harvard Classics (New York: P.F. Collier & Son, 1909-14; Bartleby.com, 2001. www.bartleby.com/19/3/).
- Goodin, R. E., 'The State as Moral Agent', in: Pettit, P., and Hamlin, A., (eds.), *The Good Polity: Normative Analysis of the State* (Oxford: Basil Blackwell, 1989), pp. 123-139.
- Goodin, R. E., 'What Is So Special about Our Fellow Countrymen?', *Ethics*, 98 (July 1989), pp. 663-686.
- Goodin, R. E., *Protecting the Vulnerable* (Chicago: University of Chicago Press, 1985).
- Goodin, R. E., and Pettit, P., 'The Possibility of Special Duties', *Canadian Journal of Philosophy*, 16 (December 1986), pp. 651-676.
- Graham, K., 'Morality, individuals and Collectives', in: J. D. G. Evans (ed.), *Moral Philosophy and Contemporary Problems* (Cambridge: Cambridge University Press, 1987), pp. 1-18.
- Graham, K., *Karl Marx, Our Contemporary* (New York: Harvester Wheatsheaf, 1992).
- Graham, K., *Practical Reasoning in a Social World: How We Act Together* (Cambridge: Cambridge University Press, 2002).
- Gray, J., *Beyond the New Right* (London & New York: Routledge, 1993).
- Green, L. C., *The Authority of the State* (Oxford: Clarendon Press, 1988).
- Grice, H. P., 'Personal Identity', in: J. Perry (ed.), *Personal Identity* (Berkeley: University of California Press, 1975), pp. 74-95.
- Hallis, F., *Corporate Personality: A Study in Jurisprudence* (Oxford: Oxford University Press, 1930).
- Hampton, J., 'Selflessness and the Loss of Self', in: Paul, E. F., Miller F. D., and Paul J., (eds.), *Altruism* (Cambridge: Cambridge University Press, 1993), pp. 135-165.
- Hampton, J., *Hobbes and the Social Contract Tradition* (Cambridge: Cambridge University Press, 1992).
- Hardimon, M. O., 'Role Obligations', *The Journal of Philosophy*, 91, Number 7 (1994), pp. 333-363.
- Hart, H. L. A., 'Are There Any Natural Rights?', in: Quinton, A., (ed.), *Political Philosophy* (London: Oxford University Press, 1967), pp. 53-66.
- Hart, H. L. A., 'Between Utility and Rights', in: *Essays in Jurisprudence and Philosophy* (Oxford: Oxford University Press, 1983), pp. 198-222.
- Hart, H. L. A., 'Definition and Theory in Jurisprudence', in: *Essays in Jurisprudence and Philosophy* (Oxford: Oxford University Press, 1983), pp. 21-48.
- Hart, H. L. A., *The Concept of Law* (Oxford: Clarendon Press, 1997), second edition.

- Hart, H. L. A., 'Postscript', in: *The Concept of Law*, P. A. Bullock and J. Raz (eds.), (Oxford: Clarendon Press, 1997), second edition.
- Hayek, F. A. von, 'Individualism: True and False', in: *Individualism and Economic Order* (London: Routledge and Kegan Paul, 1949), pp. 1-32.
- Hayek, F. A. von, *Law, Legislation and Liberty* (London: Routledge, 1982).
- Heath Wellman, C., 'Friends, Compatriots, and Special Political Obligations', *Political Theory*, Vol. 29, No. 2 (April 2001), pp. 217-236.
- Heath Wellman, C., 'Relational Facts in Political Theory: Is There Magic in the Pronoun 'My'?', *Ethics*, 110 (April 2000) pp. 537-562.
- Heath, J., 'Intergenerational Cooperation and Distributive Justice', *Canadian Journal of Philosophy*, Volume 27, Number 3 (September 1997), pp. 361-376.
- Hegel, G. F. W., *Philosophy of Right*, Knox, T. M., (ed.) (London: Oxford University Press, 1967).
- Heilbroner, R. L., 'What Has Posterity Ever Done For Me?', in: Partridge, E., (ed.) *Responsibilities to Future Generations: Environmental Ethics* (Buffalo: Prometheus Books, 1981), pp. 191-194.
- Herzog, D., 'Puzzling Through Burke', *Political Theory*, Volume 19, No. 3 (1991), pp. 336-363.
- Hirst, P., *Associative Democracy: New Forms of Economic and Social Governance* (Cambridge: Polity Press, 1994).
- Hobbes, T., *Leviathan*, Tuck, R., (ed.) (Cambridge: Cambridge University Press, 1991).
- Hollis, M., *Trust within Reason* (Cambridge: Cambridge University Press, 1998).
- Hollis, M., and Sugden, R., 'Rationality in Action', *Mind*, Volume 102, Issue 405 (January 1993), pp. 1-35.
- Holmes, S., 'Precommitment and the Paradox of Democracy', in: Elster, J., and Slagstad, R., (eds.) *Constitutionalism and Democracy* (Cambridge: Cambridge University Press, 1988), pp. 195-240.
- Holmes, S., *Passions and Constraint: On the Theory of Liberal Democracy* (Chicago: University of Chicago Press, 1995).
- Horton, J., *Political Obligation* (Basingstoke: MacMillan, 1992).
- Hubin, D. C., 'Justice and Future Generations', *Philosophy and Public Affairs*, Vol. 6, Issue 1 (1976), pp. 70-83.
- Hume, D., *A Treatise of Human Nature*, second edition by L. A. Selby-Bigge and P. H. Nidditch (Oxford: Oxford University Press, 1978).
- Hume, D., 'Of the Original Contract', in: Miller, E. F., (ed.) *Essays: Moral, Political and Literary* (Indianapolis: Liberty Fund, 1985), pp. 265-487.
- Hurley, S., *Natural Reasons: Personality and Polity* (Oxford: Oxford University Press, 1989).
- Jefferson, T., 'Debt, Taxes, Banks, and Paper', Letter to John Wayles Eppes (June 24, 1813) in: Peterson, M. D., (ed.), *Writings* (New York: The Library of America, 1984), pp. 1280-1286.
- Jefferson, T., 'Reform of the Virginia Constitution', Letter to Samuel Kercheval (July 12, 1816), in: Peterson, M. D., (ed.), *Writings* (New York: The Library of America, 1984), pp. 1395-1403.
- Jefferson, T., 'The Earth Belongs to the Living', Letter to James Madison (September 6, 1789), in: Peterson, M. D., (ed.), *Writings* (New York: The Library of America, 1984), pp. 959-964.

- Jones, P., 'Group Rights and Group Oppression', *The Journal of Political Philosophy*, Vol. 7, No 4. (1998), pp. 353-377.
- Kagan, S., *The Limits of Morality* (Oxford: Clarendon Press, 1991).
- Kant, I., 'On the Common Saying: This May be True in Theory, but it does not Apply in Practice', in: Reiss H., (ed.), *Kant: Political Writings* (Cambridge: Cambridge University Press, 1991), pp. 61-92.
- Kavka, G. S., 'Is Individual Choice Less Problematic than Collective Choice?', *Economics and Philosophy*, 7 (1991), pp. 143-165.
- Kavka, G. S., 'The Futurity Problem', in: Sikora, R. I., and Barry, B., (eds.) *Obligations to Future Generations* (Philadelphia: Temple University Press, 1978), pp. 186-203.
- Kavka, G. S., 'The Paradox of Future Individuals', *Philosophy and Public Affairs*, Volume 11 (1982), pp. 93-112.
- Kavka, G. S., 'The Toxin Puzzle', *Analysis*, 42 (1983), pp. 33-36.
- Kavka, G. S., *Hobbesian Moral and Political Theory* (Princeton: Princeton University Press, 1986).
- Korsgaard, C. M., 'Personal Identity and the Unity of Agency: A Kantian Response to Parfit', *Philosophy and Public Affairs*, Volume 18 (1989), pp. 101-132.
- Korsgaard, C. M., 'Self-Construction in the Ethics of Plato and Kant', *The Journal of Ethics* (3: 1999), pp. 1-29.
- Korsgaard, C. M., *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996).
- Kukathas, C., 'Are There Any Cultural Rights', *Political Theory*, Volume 20 (1992), pp. 105-139.
- Kutz, C., 'Acting Together', *Philosophy and Phenomenological Research*, Vol., LXI, No. 1 (July 2000), pp. 1-31.
- Kymlicka, W., *Liberalism, Community and Culture* (Oxford: Clarendon Press, 1989).
- Kymlicka, W., 'Individual and Community Rights', in: J. Baker (ed.), *Group Rights* (Toronto: University of Toronto Press, 1994), pp. 17-33.
- Laslett, P. and Fishkin, J. S., 'Introduction', in: Laslett, P., and Fishkin, J. S., (eds.) *Justice Between Age Groups and Generations* (New Haven: Yale University Press, 1992), pp. 1-32.
- Lewis, D. K., *Convention: A Philosophical Study* (Cambridge: Harvard University Press, 1969).
- List C., and Pettit, P., 'Aggregating Sets of Judgments: An Impossibility Result', *Economics and Philosophy*, 18 (2002), pp. 89-110.
- Locke, J., *An Essay Concerning Human Understanding*, P. H. Nidditch (ed.) (Oxford: Oxford University Press, 1975).
- Locke, J., *Second Treatises of Government*, edited by C. B. Macpherson (Indianapolis: Hackett Publishing Company, 1980).
- Lowenthal, D., *The Past is a Foreign Country* (Cambridge: Cambridge University Press, 1985).
- Lukes, S., *Individualism* (Oxford: Basil Blackwell, 1985).
- Lyons, D., 'The Moral Opacity of Utilitarianism', in: Hooker, B., Mason, E., and Miller, D. E., (eds.) *Morality, Rules, and Consequences: A Critical Reader* (Edinburgh: Edinburgh University Press, 2000), pp. 105-120.
- Machen, A. W., 'Corporate Personality', *Harvard Law Review*, Volume XXIV, No. 4 (February, 1911), pp. 253-267.

- MacIntyre, A., *After Virtue: A Study in Moral Theory* (London: Duckworth, 2001).
- Macklin, R., 'Can Future Generations Correctly Be Said to Have Rights', in: Partridge, E., (ed.) *Responsibilities to Future Generations: Environmental Ethics* (Buffalo: Prometheus Books, 1981), pp. 151-155.
- MacLean, D., 'A Moral Requirement of Energy Policies', in: Maclean, D., and Brown, P. G., (eds.), *Energy and The Future* (Totowa: Rowman and Littlefield, 1983), pp. 180-197.
- Maitland, F. W., 'Introduction', in: O. Gierke, *Political Theories of the Middle Age*, F. W. Maitland (ed.), (Cambridge: Cambridge University Press, 1927).
- Maitland, F. W., 'Moral Personality and Legal Personality', in: H. A. L. Fisher (ed.), *The Collected Papers of Frederick William Maitland*, Volume III (Cambridge: Cambridge University Press, 1911), pp. 304-320.
- March, J. G., and Olsen, J. P., *Rediscovering Institutions: The Organizational Basis of Politics* (New York: The Free Press, 1989).
- Martin, R., and Barresi, J., *Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century* (London: Routledge, 1999).
- Mason, A., 'Special Obligations to Compatriots', *Ethics*, 107 (April 1997), pp. 427-447.
- Mason, A., *Community, Solidarity and Belonging* (Cambridge: Cambridge University Press, 2000).
- May, L., and Hoffman, S., (eds.), *Collective Responsibility: Five Decades of Debate in Theoretical and Applied Ethics* (Savage: Rowman & Littlefield Publishers, Inc., 1991).
- May, L., *The Morality of Groups: Collective Responsibility, Group-based Harm, and Corporate Rights* (Notre Dame: University of Notre Dame, 1987).
- McClennen, E. F., *Rationality and Dynamic Choice: Foundational Explorations* (Cambridge: Cambridge University Press, 1990).
- McClennen, E. F., 'Rationality and Rules', in: Danielson, P. A., (ed.) *Modelling Rationality, Morality, and Evolution* (Oxford: Oxford University Press, 1998), pp. 13-40.
- McDowell, J., 'Reductionism and the First Person', in: J. Dancy (ed.), *Reading Parfit* (Oxford: Blackwell Publishers, 1997), pp. 230-250.
- McMahan, J., 'Nationalism and the Demands of Impartiality', in: McKim, R., and McMahan, J., (eds.), *The Morality of Nationalism* (New York: Oxford University Press, 1997), pp. 107-138.
- McMahon, C., *Authority and Democracy: A General Theory of Government and Management* (Princeton: Princeton University Press, 1994).
- Miller, D., *On Nationality* (Oxford: Oxford University Press).
- Miller, D., 'Nationality: Some Replies', *The Journal of Applied Philosophy*, 14 (1997), pp. 69-82.
- Miller, D., *Citizenship and National Identity* (Cambridge: Polity Press, 2000).
- Miller, D., *Market, State and Community* (Oxford: Clarendon Press, 1989).
- Millgram, E., 'Incommensurability and Practical Reasoning', in: Chang, R., (ed.) *Incommensurability, Incomparability, and Practical Reason* (Cambridge: Harvard University Press, 1997), pp. 151-169.
- Mulgan, T., 'The Place of the Dead in Liberal Political Philosophy', *The Journal of Political Philosophy*, Vol. 7 (1999), pp. 52-70.
- Nagel, T., *The Possibility of Altruism* (Princeton: Princeton University Press, 1970).
- Nagel, T., *Equality and Partiality* (New York: Oxford University Press, 1991).

- Narveson, J., 'Future People and Us', in: Sikora, R. I., and Barry, B., (eds.), *Obligations to Future Generations* (Cambridge: The White Horse Press, 1996), pp. 38-60.
- Nicholls, D., *The Pluralist State: The Political Ideas of J. N. Figgis and his Contemporaries*, second edition (Basingstoke: St. Martin's Press, 1994).
- Nietzsche, F., *Twilight of the Idols: Or How One Philosophizes With a Hammer*, in: Kaufmann, W., (ed.), *The Portable Nietzsche* (Harmondsworth: Penguin, 1984).
- Nozick, R., *Anarchy, State, and Utopia* (New York: Basic Books, 1974).
- Oakeshott, M., *On Human Conduct* (Oxford: Oxford University Press, 1975).
- O'Neill, J., *Ecology, Policy and Politics: Human Well-Being and the Natural World* (London: Routledge, 1993).
- Oldenquist, A., 'Loyalties', *The Journal of Philosophy*, Vol. LXXIX, No. 4 (April 1984), pp. 173-193.
- Paden, R., 'Reciprocity and Intergenerational Justice', *Public Affairs Quarterly*, Vol. 10, Number 3 (July, 1996), pp. 246-266.
- Page, E., 'Intergenerational Justice and Climate Change', *Political Studies*, XLVII (1999), pp. 53-66.
- Paine, T., *Rights of Man*, Foner, P. S., (ed.) (Secaucus: The Citadel Press, 1974).
- Parekh, B., 'A Misconceived Discourse on Political Obligation', *Political Studies*, XLI (1993), pp. 236-251.
- Parfit, D., 'Personal Identity', *The Philosophical Review*, Vol. 80, Issue 1 (January, 1971), pp. 3-27.
- Parfit, D., 'On 'The Importance of Self-Identity'', *The Journal of Philosophy*, Vol. 68, Issue 20 (October 1971), pp. 683-690.
- Parfit, D., 'Later Selves and Moral Principles', in: Montefiore, A., (ed.), *Philosophy and Personal Relations: An Anglo-French Study* (London: Routledge and Kegan Paul, 1973), pp. 137-169;
- Parfit, D., 'Future Generations: Further Problems', *Philosophy and Public Affairs*, Volume 11 (1982), pp. 113-172.
- Parfit, D., 'Personal Identity and Rationality', *Synthese*, Vol. 53, No.2 (November 1982), pp. 227-241.
- Parfit, D., 'Comments', *Ethics*, Volume 96, No. 4. (July 1986), pp. 832-872.
- Parfit, D., *Reasons and Persons*, third and revised edition (Oxford: Clarendon Press, 1987).
- Parfit, D., 'Divided Minds and the Nature of Persons', in: C. Blakemore and S. Greenfield (eds.), *Mindwaves: Thoughts on Intelligence, Identity and Consciousness* (Oxford: Basil Blackwell, 1987), pp. 19-28.
- Parfit, D., 'The Unimportance of Identity', in: H. Harris (ed.), *Identity: Essays Based on Herbert Spencer Lectures Given in the University of Oxford* (Oxford: Clarendon press, 1995), pp. 13-45.
- Partridge, E., 'Why Care About the Future?', in: Partridge, E., (ed.) *Responsibilities to Future Generations: Environmental Ethics* (Buffalo: Prometheus Books, 1981), pp. 203-220.
- Perry, J., 'Personal Identity, Memory, and the Problem of Circularity', in: Perry, J., (ed.), *Personal Identity* (Berkeley: University of California Press, 1975), pp. 135-155.
- Perry, J., 'The Problem of Personal Identity', in: Perry, J., (ed.), *Personal Identity* (Berkeley: University of California Press, 1975), pp. 3-30.

- Pettit, P., *The Common Mind: An Essay on Psychology, Society and Politics* (Oxford: Oxford University Press, 1993).
- Pettit, P., 'Akrasia, Collective and Individual', *Australia National University Working Papers*, <http://socpol.anu.edu.au>, 2001.
- Pettit, P., 'Deliberative Democracy and the Discursive Dilemma', *Australia National University Working Papers*, <http://socpol.anu.edu.au>, 2001.
- Pettit, P., 'Groups with Minds of their Own', *Australia National University Working Papers*, <http://socpol.anu.edu.au>, 2001.
- Pettit, P., *A Theory of Freedom* (Cambridge: Polity Press, 2001).
- Plato, *The Republic*, Lee, D., (ed.), second edition (London: Penguin Books, 1987).
- Popper, K. R., *The Open Society and its Enemies*, Volume 2, (London: Routledge, 1945).
- Postema, G. J., 'Morality in the First Person Plural', *Law and Philosophy*, 14 (1995), pp. 35-64.
- Quinton, A., 'The Soul', in: J. Perry (ed.), *Personal Identity* (Berkeley: University of California Press, 1975), pp. 53-72.
- Quinton, A., 'Social Objects', *The Aristotelian Society*, Volume LXXVI (1975/76), pp. 1-27.
- Rawls, J., *A Theory of Justice* (Oxford: Oxford University Press, 1992).
- Rawls, J., 'Justice as Fairness: Political not Metaphysical', *Philosophy and Public Affairs*, Vol. 14 (1985), pp. 223-251.
- Rawls, J., 'The Independence of Moral Theory', in: S. Freeman (ed.), *John Rawls: Collected Essays* (Cambridge: Harvard University Press, 1999), pp. 286-302.
- Rawls, J., *Political Liberalism* (New York: Columbia University Press, 1993).
- Rawls, J., *Justice as Fairness: A Restatement*, Kelly, E., (ed.) (Cambridge MA: The Belknap Press of Harvard University Press, 2000).
- Raz, J., 'Liberating Duties', *Law and Philosophy*, Vol. 8 (1989), pp. 3-21.
- Réaume, D., 'The Group Right to Linguistic Security: Whose Right, What Duties', in: Baker, J., *Group Rights* (Toronto: University of Toronto Press, 1994), pp. 118-141.
- Reid, T., 'Of Identity', in: J. Perry (ed.), *Personal Identity* (Berkeley: University of California Press, 1975), pp. 107-112.
- Reid, T., 'Of Mr. Locke's Account of our Personal Identity', in: J. Perry (ed.), *Personal Identity* (Berkeley: University of California Press, 1975), pp. 113-118.
- Richards, D. A. J., 'Contractarian Theory and Intergenerational Justice', in: Maclean, D., and Brown, P. G., (eds.), *Energy and The Future* (Totowa: Rowman and Littlefield, 1983), pp. 131-150.
- Rousseau, J. J., *The Social Contract*, in: Cole, G. D. H., (ed.) *The Social Contract and Discourses* (London: Everyman, 1997).
- Rovane, C., 'Self-Reference: The Radicalization of Locke', *The Journal of Philosophy*, Vol. 90, Issue 2, (1993), pp. 73-97.
- Rovane, C., *The Bounds of Agency: An Essay in Revisionary Metaphysics* (Princeton: Princeton University Press, 1998).
- Ruben, D. H., *The Metaphysics of the Social World* (London: Routledge and Kegan Paul, 1985).
- Runciman, D., *Pluralism and the Personality of the State* (Cambridge: Cambridge University Press, 1997).

- Runciman, D., 'What Kind of Person is Hobbes's State? A Reply to Skinner', *The Journal of Political Philosophy*, Volume 8, number 2 (2000), pp. 268-278.
- Sandel, M. J., *Liberalism and the Limits of Justice* (Cambridge: Cambridge University Press, 1982).
- Sauvé, K., 'Gauthier, Property Rights, and Future Generations', *Canadian Journal of Philosophy*, Volume 25, Number 2 (1995), pp. 163-176.
- Scanlon, T. M., 'Contractualism and Utilitarianism', in: Sen, A., and Williams, B., (ed.), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1994), pp. 103-128.
- Scanlon, T. M., *What We Owe to Each Other* (Cambridge: The Belknap Press, 1998).
- Scheffler, S., 'Ethics, Personal Identity, and Ideals of the Person', *Canadian Journal of Philosophy*, Vol. XII, No. 2 (June 1982), pp. 229-246.
- Scheffler, S., 'Families, Nations, and Strangers', in: *Boundaries and Allegiances: Problems of Justice and Responsibility in Liberal Thought* (Oxford: Oxford University Press, 2000), pp. 48-65.
- Scheffler, S., 'Liberalism, Nationalism, and Egalitarianism', in: *Boundaries and Allegiances: Problems of Justice and Responsibility in Liberal Thought* (Oxford: Oxford University Press, 2000), pp. 66-81.
- Scheffler, S., 'The Conflict between Justice and Responsibility', in: *Boundaries and Allegiances: Problems of Justice and Responsibility in Liberal Thought* (Oxford: Oxford University Press, 2000), pp. 82-96.
- Scheffler, S., in 'Relationships and Responsibilities', in: *Boundaries and Allegiances: Problems of Justice and Responsibility in Liberal Thought* (Oxford: Oxford University Press, 2000), pp. 97-110.
- Schultz, B., 'Persons, Selves, and Utilitarianism', *Ethics*, Volume 96, No. 4. (July, 1986), pp. 721-745.
- Schwartz, T., 'Obligations to Posterity', in: Sikora, R. I., and Barry, B., (eds.), *Obligations to Future Generations* (Philadelphia: Temple University Press, 1978), pp. 3-13.
- Scruton, R., 'Corporate Persons', *Proceedings of the Aristotelian Society*, 63 (1989), pp. 239-266.
- Searle, J. R., 'Collective Intentions and Actions', in: Cohen, P. R., Morgan, J., and Pollack, M. E., (eds.), *Intentions in Communication* (Cambridge M.A.: The MIT Press, 1990), pp. 401-415.
- Searle, J. R., *The Construction of Social Reality* (London: Penguin Books, 1996).
- Sen, A., 'Utilitarianism and Welfarism', *The Journal of Philosophy*, Vol. 76, No. 9. (Sep., 1979), pp. 463-489.
- Sen, A., and Williams, B., 'Introduction: Utilitarianism and Beyond', in: Sen, A., and Williams, B., (eds.) *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1994), pp. 15-16.
- Shapiro, S. J., 'Ulysses Rebound', *Economics and Philosophy*, 18 (2002), pp. 157-182.
- Shoemaker S., 'Persons and their Past', *American Philosophical Quarterly* 7 (1970), pp. 269-285.
- Shoemaker S., 'A Materialist Account', in: Shoemaker, S., and Swinburne, R., (eds.), *Personal Identity* (Oxford: Basil Blackwell, 1984), pp. 67-132.
- Simmons, A. J., 'Associative Political Obligations', *Ethics*, 106 (January 1996), pp. 247-273.

- Skinner, Q., 'Hobbes and the Purely Artificial Person of the State', *The Journal of Political Philosophy*, Volume 7, Number 1 (1999), pp. 1-29.
- Slote, M., 'Goods and Lives', in: *Goods and Virtues* (Oxford: Clarendon Press, 1983), pp. 9-37.
- Smart J. J. C., and Williams, B., *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1973).
- Steiner, H., 'The Rights of Future Generations', in: Maclean, D., and Brown, P. G., (eds.), *Energy and The Future* (Totowa: Rowman and Littlefield, 1983), pp. 151-165.
- Stoljar, S. J., *Groups and Entities: An Inquiry into Corporate Theory* (Canberra: Australia National University Press, 1973).
- Stone, J., 'Parfit and the Buddha: Why There Are No People', *Philosophy and Phenomenological Research*, Volume 48, No. 3. (1988), pp. 519-532.
- Sugden, R., 'Thinking as a Team', in: Paul, E. F., Miller, F. D., and Paul, J., (eds.), *Altruism* (Cambridge: Cambridge University Press, 1993), pp. 69-89.
- Sugden, R., 'Rational Co-ordination', in: Farina, F., Hahn, F., and Vannucci, S., (eds.) *Ethics, Rationality, and Economic Behaviour* (Oxford: Clarendon Press, 1996), pp. 244-262.
- Sugden, R., 'Team Preferences', *Economics and Philosophy*, 16 (2000), pp. 175-204.
- Sugden, R., 'Beyond Sympathy and Empathy: Adam Smith's Concept of Fellow-Feeling', *Economics and Philosophy*, 18 (2002), pp. 63-87.
- Swinburne, R., 'Personal Identity: the Dualist Theory', in: Shoemaker, S., and Swinburne, R., (eds.), *Personal Identity* (Oxford: Basil Blackwell, 1984), pp. 1-66.
- Tamir, Y., *Liberal Nationalism* (Princeton: Princeton University Press, 1993).
- Taylor, C., *Sources of the Self: The Making of the Modern Identity* (Cambridge: Cambridge University Press, 1989).
- Taylor, C., 'Lichtung or Lebensform' in: *Philosophical Arguments* (Cambridge MA: Harvard University Press, 1995), pp. 61-78.
- Taylor, C., 'The Importance of Herder', in: *Philosophical Arguments* (Cambridge MA: Harvard University Press, 1995), pp. 79-99.
- Taylor, C., 'Irreducibly Social Goods', in: *Philosophical Arguments* (Cambridge MA: Harvard University Press, 1995), pp. 127-145.
- Taylor, C., 'Cross-Purposes: The Liberal-Communitarian Debate', in: *Philosophical Arguments* (Cambridge MA: Harvard University Press, 1995), pp. 181-203.
- Temkin, L. S., 'Justice and Equality: Some Questions About Scope', *Social Philosophy and Policy*, Volume 12, Number 2 (Summer 1995), pp. 72-104.
- Teubner, G., 'Enterprise Corporatism: New Industrial Policy and the 'Essence' of the Legal Person', in: Wheeler, S., (ed.) *A Reader on the Law of Business Enterprise* (Oxford: Oxford University Press, 1994), pp. 51-79.
- Thompson, T. T., 'Are We Obligated to Future Others?', in: Partridge, E., (ed.), *Responsibilities to Future Generations* (Buffalo/New York: Prometheus Books, 1981), pp. 195-202.
- Tönnies, F., *Community and Civil Society*, Harris, J., (ed.), (Cambridge: Cambridge University Press, 2001).
- Treaty of Utrecht (1713), Cited in: A. Heath, 'Research Roundup: The Assault on Gibraltar', *The European Journal*, Vol. 9, No. 3 (2001), pp. 8-9, p. 9.

- Tuomela, R., 'We Will Do It: An Analysis of Group-Intentions', *Philosophy and Phenomenological Research*, Volume 51, Issue 2 (June 1991), pp. 249-277.
- Tuomela, R., *Cooperation: A Philosophical Study* (Dordrecht: Kluwer Academic Publishers, 2000).
- Udehn, L., *Methodological Individualism: Background, History and Meaning* (London: Routledge, 2001).
- United Nations, *Report of the United Nations Conference on Environment and Development*, at <http://www.un.org/documents/ga/conf151> (1992).
- Velleman, J. D., 'What Happens When Someone Acts', *Mind*, Vol. 101, Issue 403 (1992), pp. 461-481.
- Velleman, J. D., 'Well-being and Time', in: *The Possibility of Practical Reason*, (Oxford: Oxford University Press, 2000), pp. 170-199.
- Velleman, J. D., 'The Possibility of Practical Reason', in: *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000), pp. 170-199.
- Velleman, J. D., 'How to Share an Intention', in: *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000), pp. 200-220.
- Waddell Ekstrom, L., 'A Coherence Theory of Autonomy', *Philosophy and Phenomenological Research*, Volume 53, Issue 3 (1993), pp. 599-616.
- Waldron, J., 'Communal Goods as Human Rights', in: *Liberal Rights* (Cambridge: Cambridge University Press, 1993) pp. 339-369.
- Waldron, J., 'Special Ties and Natural Duties', *Philosophy and Public Affairs*, Volume 22, Issue 1 (Winter, 1993), pp. 3-30.
- Watson, G., 'Free Agency', in: Watson, G., (ed.) *Free Will* (Oxford: Oxford University Press, 1982), pp. 96-110.
- White, H., *Meta-history: The Historical Imagination in Nineteenth-Century Europe* (Baltimore: Johns Hopkins University Press, 1973).
- White, H., 'Historical Emplotment and the Problem of Truth', in: Roberts, G., (ed.), *The History and Narrative Reader* (London: Routledge, 2001), pp. 375-389.
- Whiting, J., 'Friends and Future Selves', *Philosophical Review*, Volume 95, Issue 4 (1986), pp. 547-580.
- Williams, B., 'Persons, Character and Morality', in Rorty, A. O., (ed.), *The Identities of Persons* (Berkeley: University of California Press, 1976), pp. 197-216.
- Williams, B., 'The Self and the Future', in: Perry, J., (ed.), *Personal Identity* (Berkeley: University of California Press, 1975), pp 179-198.
- Williams, B., 'Internal and External Reasons', in: *Moral Luck: Philosophical Papers 1973-1980* (Cambridge: Cambridge University Press, 1981), pp.101-113.
- Williams, B., *Ethics and the Limits of Philosophy* (London: Fontana Press, 1994).
- Wolf, C., 'Contemporary Property Rights, Lockean Provisos, and the Interests of Future Generations', *Ethics*, 105 (July 1995), pp. 791-818.
- Wolf, S. R., 'Self-Interest and Interest in Selves', *Ethics*, 96 (1986), pp. 704-720.
- Wolf, S. R., *Freedom within Reason* (New York: Oxford University Press, 1990).
- Wolf, S. R., 'Happiness and Meaning: Two Aspects of the Good Life', *Social Philosophy and Policy*, Volume 14, Number 1 (1997), pp. 207-225.
- Wolff, M., 'On the Nature of Legal Persons', *Law Quarterly Review*, Volume LIV (1938), pp. 494-521.