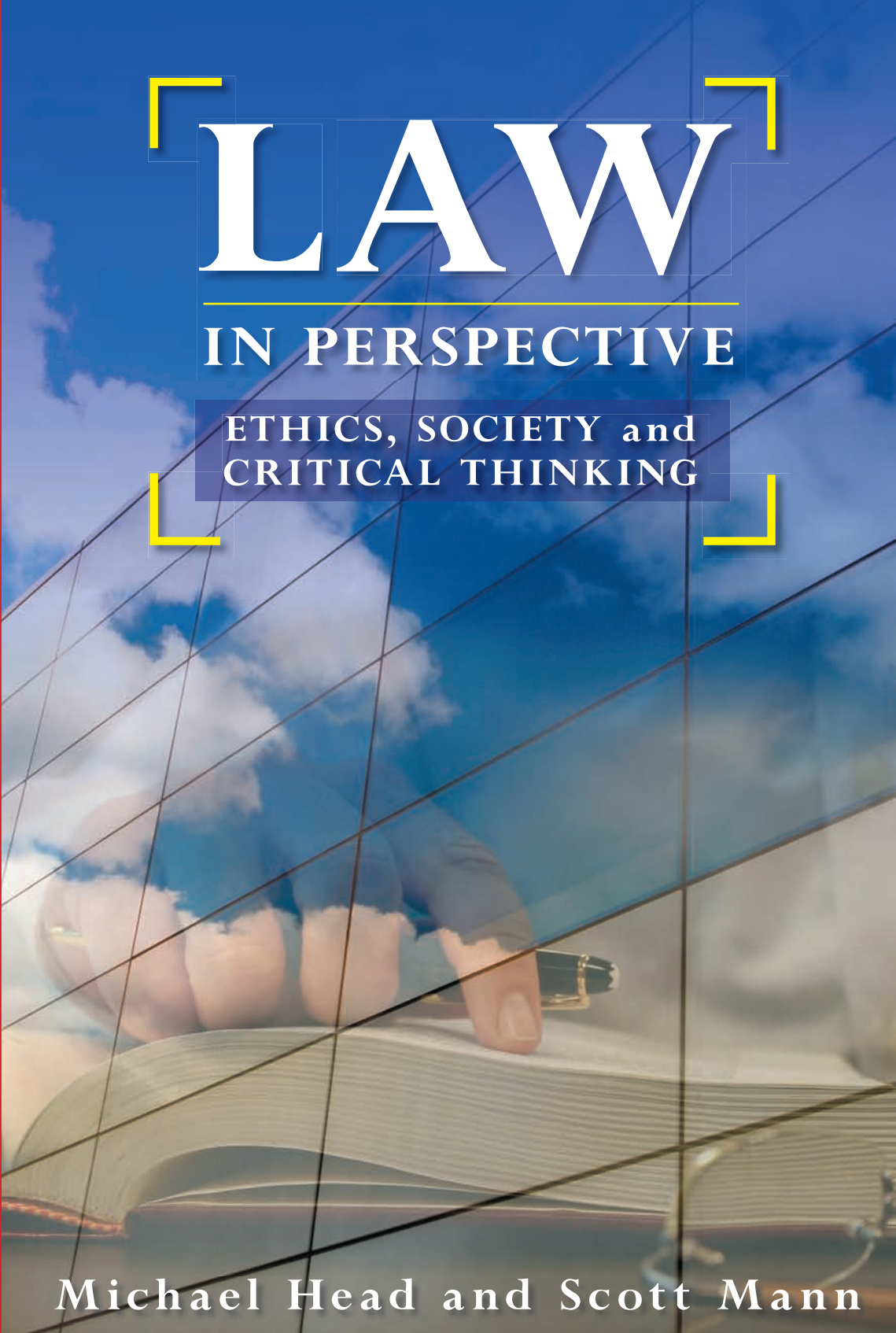

LAW
IN PERSPECTIVE
ETHICS, SOCIETY and CRITICAL THINKING
Michael Head and Scott Mann

LAW IN PERSPECTIVE

Having previously taught law at Columbia, LaTrobe, Adelaide and ANU in the 1970s, DR MICHAEL HEAD returned to academic life in 1999 and teaches at the University of Western Sydney. The author of *Administrative Law: Context and Critique*, Head also publishes articles regularly in the fields of socialist legal theory, civil liberties and refugee law. He is editor of *University of Western Sydney Law Review*.

SCOTT MANN has taught philosophy, psychology and critical social theory in the Universities of Sussex, Sydney and New South Wales. He has many years experience of teaching in adult education with the WEA in Sydney. He has published books on psychoanalysis and society, the sociology of religion, and business ethics. Prior to joining the University of Western Sydney, where he now teaches, Dr Mann was Director of the Centre for Liberal and General Studies at the University of New South Wales.

LAW
IN PERSPECTIVE

ETHICS, SOCIETY AND CRITICAL THINKING

MICHAEL HEAD AND SCOTT MANN

UNSW PRESS

A UNSW Press book

Published by
University of New South Wales Press Ltd
University of New South Wales
Sydney NSW 2052
AUSTRALIA
www.unswpress.com.au

First published 2005

National Library of Australia
Cataloguing-in-Publication entry

Head, Michael, 1952– .
Law in perspective: ethics, society and critical thinking

Includes index.
For tertiary students.
ISBN 0 86840 834 4.

1. Sociological jurisprudence. 2. Law and ethics. 3. Law.
I. Mann, Scott. II. Title.

340.115

Design Di Quick
Print Ligare

CONTENTS

PREFACE

For their unique role in inspiring and assisting us to prepare what we hope will prove to be a pioneering text, our first thanks go to the students we have taught over the years in Law Foundation at the University of Western Sydney. Their questions, comments and criticisms of our efforts to place law in perspective have been invaluable.

Every year, we set out to prove to our students that our course is the most important they will do in studying law-related subjects. All too often, through no fault of their own, students have been initially unfamiliar with many of the historical and theoretical issues posed. But they have enthusiastically taken up the task of acquainting themselves with the contributions made over centuries by great thinkers in logic, science, ethics, legal theory and social sciences. For us, every year has been a rewarding experience, as we have developed and enriched the subject, and we trust that this is reflected in the book.

Thanks must also go to our colleagues at UWS, notably Carolyn and Razeen Sappideen, for their support and encouragement. We are also grateful to everyone at UNSW Press for their fine work in bringing this volume to fruition.

Michael Head owes a particular debt to his partner, Mary, and children, Tom, Daniel and Kathleen, for their love and patience throughout the long hours and labours devoted to this work

Scott Mann would also like to thank Kay, Jocelyn and Claire for lots of moral support.

Michael Head and Scott Mann

INTRODUCTION

The principal aim of this book is to encourage critical, responsible and creative thinking about law as a system of ideas and as a social institution. It aims also to encourage exploration of the interrelations between legal methods, ideas and practices and those of other disciplines and institutions with which law continuously interacts.

To this end, the book focuses upon a range of powerful critical thinking tools, in the form of ideas and techniques drawn from logic, science, ethics and political and social theory. Effective application of these tools allows for an appreciation of law in its historical, philosophical, economic, political and social context. It provides a foundation for critical assessment of the value and significance of particular legal principles and practices, and for principled proposals for future legal development and reform.

In later chapters, the book demonstrates such tools in action, developing critical analyses of the role of law in relation to a range of contemporary social issues. Problems with existing legal interventions are highlighted and possible alternative approaches explored.

The book has been designed, first and foremost, as an introductory textbook, following on from a basic introduction to the common law system. But no detailed or specialised knowledge of the legal system is required in order to understand and benefit from it.

At the same time, we believe that this material will be of broad interest, not only to law students, but also to legal academics and practitioners, social policy and welfare academics, students and practitioners, professional and business people and all active and concerned citizens.

It is all too easy for starting law students to be overwhelmed by the sheer mass of data they are expected to master in a relatively short space of time. Some law teachers and practitioners argue that this leaves no time for critical analysis. And some maintain that 'the basics' have to be mastered before meaningful critical analysis is possible.

The fact that some teachers have equated a 'critical and interdisciplinary approach' with woolly ramblings or faddish postmodern deconstructions lacking in all substantive content has appeared to provide further support for a strongly disciplinary and black-letter approach.

But the dangers of such positivism are substantial. To discourage critical thinking from the start is to encourage an uncritical acceptance of what are, in some cases, logically and morally unacceptable ideas, practices and consequences. To discourage interdisciplinarity is to encourage a view of law as a self-subsistent system of ideas and values closed off from all broader social significance, impact and responsibility, and shut off from effective comprehension, participation and critical assessment by all but qualified legal specialists.

In contrast to such a positivist perspective, we believe that it is crucially important for students to master basic tools of critical thinking and analysis at the earliest possible stage of their studies. This helps them to clarify, develop and apply their own values and priorities. Through empowering students, as active participants in the learning process, it makes the whole process a great deal more exciting and enjoyable.

Through challenging and questioning, rather than merely accepting, through exploring the interrelations of law with other disciplines, ideas and social practices, students acquire an altogether deeper, more grounded, more responsible and multidimensional understanding of legal ideas, methods, practices and institutions. They are in a better position to understand, anticipate and evaluate new legal developments and trends in an epoch of far-reaching change on a global scale.

Over recent years, courses on law and its social context have become cornerstones of undergraduate law school degrees. The Pearce report, an assessment of legal education released in 1987, criticised Australian law schools for neglecting the critical and theoretical dimensions of law required for a flourishing intellectual academic legal culture and an understanding of what role law and lawyers should play in society.[1]

Most university law schools now insist that aspiring lawyers and others seeking law-related careers have at least some exposure to the sorts of questions canvassed in this book. A 1994 review of the impact of the Pearce report found that most law schools attached 'considerable importance' to the critical examination of legal issues in their social context, while noting that the expression and emphasis of such issues varied.[2]

Almost all students now do their law degree as a joint degree or as a second degree, rather than as a purely professional training course. However, the benefits of this paradigm shift in legal education have been diminished by the tendency of many students, acting under increasing economic pressure, to make their other degree a relatively narrow-focused business or commerce degree. These combinations can mean that students have little exposure to the methodologies and critiques of the social – or natural – sciences.

The need to counteract this tendency is amplified by the apprehension that, through no fault of their own, few students come to law school with any adequate grounding in intellectual history, let alone legal and constitutional history. The study of history in secondary schools has often been relegated in favour of more vocationally oriented subjects. To take just one example, in our several years of experience in teaching Law Foundation, the introductory law, theory and society course at the University of Western Sydney, we have yet to find a new school-leaver student acquainted with the name John Locke. Locke's writings are central to understanding two pivotal features of the legal system: the separation of powers doctrine and the inviolability of private property.

Without historical and theoretical background, it is difficult to recognise that law is an instrument of social regulation that has been fashioned, and continues to be shaped, by deep-rooted economic and political factors. The existing socio-economic structure of society and its corresponding legal framework can therefore appear to have a permanence, inevitability and even naturalness that belie the historical record of convulsive changes (such as the fall of the Roman Empire, the English, American and French Revolutions of the 17th and 18th centuries and the emergence of communism and fascism in the 20th century).

Even within the realm of the existing western legal system, it is impossible to grasp the content and significance of pervasive doctrines such as 'the separation of powers', 'natural justice' habeas corpus and

the presumption of innocence without some knowledge of the historical battles that produced them.

On one level, this book seeks to fill a yawning gap – the current lack of a text or published set of readings for introductory (and compulsory) Australian university and TAFE courses on Law and Society (for example, Law Foundation at UWS). Until now, these subjects have rested upon internally produced volumes of reproduced materials, and have lacked a core or synthesising volume. We hope that the present work will remedy the situation, perhaps augmented by other selected materials.

More broadly, by providing substantial critiques of existing works on law and its relationship to logic, science, social theory and contemporary social issues, we seek to make a contribution to scholarly and civic debate. Our approaches to the subject matter are not purely pedagogical. We believe it is necessary to combine logical, scientific and intellectual rigor with a humane, ethical and progressive approach.

In a number of chapters, we advance alternative perspectives that challenge the general adherence to market-driven and nation-state-based viewpoints. Whether or not our readers – they may be students or people working at the legal and social coalfaces – agree with our analyses, we are confident that our research and insights will provoke thought, discussion and debate.

Method of approach

The book is divided into three main sections. The first section, Law, Logic and Science, explores the relations between these three disciplines. It shows how a range of key ideas, insights and techniques drawn from logic and the philosophy of science can be of significant practical benefit to law students and legal practitioners, as well as those in other disciplines and other professions. A basic understanding of such principles can help in the organisation and analysis of complex information, in more effective reasoning and argument and particularly in interdisciplinary research and communication.

Chapters 1 and 2 provide a concise introduction to logical reasoning, including the role of conditional statements, processes of deductive and inductive reasoning and argument, and explanation and analysis of extended logical arguments. They explore the central role of logical reasoning in the historical development and day-to-day operation of the common law.

Chapters 3 and 4 provide a systematic examination of common errors of reasoning – or fallacies – including formal fallacies, fallacies of relevance, inductive fallacies and others. They focus particularly upon fallacies most common in legal contexts, including the prosecutors' fallacy, and misinterpretations of DNA evidence. They explore some of the dire social consequences of fallacious reasoning in law.

Chapters 5, 6 and 7 provide an introduction to general features of scientific method, focusing particularly upon the means for establishing the existence of causal relationships. The reader is introduced to simple but powerful statistical techniques, of increasing importance in many different areas, and of particular significance in establishing causal relationships in populations. Also considered are the precautionary principle and fallacies of hypothesis testing.

The second section of the book, Law, Ethics and Social Theory, explores some of the principal theoretical approaches to the nature and social role of law. It explains and critically analyses a range of ideas and models of the social and ethical foundations and consequences of current legal institutions and practices.

Chapters 8 and 9 provide an introduction to ethical ideas and ethical reasoning, including basic principles of metaethics and normative ethics and their application to law. These chapters explore the strengths and weaknesses of utilitarian theory, and of the ethical theories of Kant and Ross, as well as considering religiously based ethical ideas. Other topics covered include ethics and free will, guilt and conscience, rights and freedoms and equality and justice.

Chapter 10 sketches the two most prominent philosophical theories of law – natural law and legal positivism – putting them in their historical perspective. It outlines the contradictory evolution of natural law, noting its classical, revolutionary, reactionary, conservative and 'human rights' phases, as well as the dominant positivist backlash in the 19th and 20th centuries. It critically appraises the post–World War II debates provoked by the crimes of fascism and Stalinism, followed by the revival of natural law under the rubric of human rights. It also looks briefly at contemporary debates about distributive justice, centred around the work of Rawls, Nozick and Posner.

Chapter 11, on liberalism and law, traces the currently dominant forms of liberal thinking, and their differing approaches to the nature and social role of law. Neo-liberal and social-liberal ideas are contrasted with socialist thinking.

Chapter 12, on Marxism and law, outlines the classical Marxist propositions on law and the state, and tackles some common misconceptions. The collapse of the Soviet Union and its satellite states has led many to conclude that Marxism is dead and communism is no longer an alternative worth considering. In our view, this confuses Stalinism with communism. In any case, no study of law can ignore the considerable historical influence of Marxism over the past 150 years. Some of the issues canvassed are revolution, participatory democracy, the possibility of law 'withering away' and the rich experiences of the early (pre-Stalinist) years of the Russian Revolution.

Chapters 13 and 14, on law and economics, consider and question the economic assumptions of neo-liberalism, particularly ideas of the superior efficiency of market relations compared to other systems of organisation of production and distribution, and the need for law to support and/or mimic such market efficiency.

Chapters 15 and 16 focus upon questions of distributive justice – of what would constitute a fair and just system of distribution of property. Here again the major focus is upon capitalist market relations, with critical consideration of the claim that markets deliver effective equality of reward; that individuals are rewarded in proportion to their valuable social contribution. Topics covered include fair wages, inheritance and unearned income.

The final section of the book, Law and Contemporary Social Problems, asks what role law plays in alleviating or exacerbating some modern social conflicts. It also draws attention to the need to consider the social impact of the legal system, and the stark reality of some of its institutions, such as prisons.

Chapter 17, on tort law 'reform', probes what lies behind the drive to limit corporate civil liability to consumers, clients and employees. It does so in the context of overcoming the root causes of personal injuries. The chapter advances an alternative based on social compensation.

Chapter 18 looks into issues of free will and criminal culpability. It focuses upon the defence of duress, and considers the class basis of the criminal justice system. Logical and practical problems of the current system are identified and explained.

Chapter 19, on crime and punishment, considers the nature, rationales and objectives of punishment in the criminal law. Prison theory and practice are questioned and compared, with a view to possible alternatives.

Among the most contentious and far-reaching issues of the opening years of the 21st century are terrorism, refugees and the erosion of democratic rights. These are the subjects of the final chapters.

Chapter 20, on terrorism and democratic rights, examines how the so-called war on terrorism has been utilised as a pretext for the dismantling of basic democratic rights. It questions the introduction of sweeping definitions of terrorism, detention without trial, powers to ban political groups and measures to allow the armed forces to be mobilised to put down domestic unrest. The chapter reviews the military call-out legislation of 2000 and the anti-terrorism laws of 2002–03 in some detail.

The concluding chapter 21, on refugees and the nation-state, documents the growing global refugee crisis and outlines the curtailing of refugee rights and the need for a new global perspective. Among the topics are the *Tampa* case, the banishing of asylum seekers and the underlying flaws of the Refugee Convention.

Each chapter includes some questions for discussion, along with some extra readings that complement or extend the discussion in the text.

While this volume has been a collaborative effort, Scott Mann has been primarily responsible for chapters 1 to 9, 11, 13 to 17 and 19 and Michael Head for chapters 10, 12, 20 and 21. Chapter 18 was a joint project between Scott and Mouaid Al Qudah. We welcome any comments or suggestions.

SECTION ONE

Logic, science and law

Chapter 1

BASIC CONCEPTS OF LOGICAL REASONING*

Logic

Logic is 'the science that evaluates arguments'. As Patrick Hurley says in his *Concise Introduction to Logic*, 'the aim of logic is to develop a system of methods and principles that we may use as criteria for evaluating the arguments of others and as guides in constructing arguments of our own'.[1]

An argument, as the term is used in logic, refers to a group of statements, one or more of which – the premises – 'are claimed to provide support for, or reason to believe, one of the others – the conclusion'.[2] A statement is a representation that is either true or false. Typically, this will be a declarative sentence, but statements can be made in many different ways, including the use of pictures and gestures.

In Hurley's words: 'The premises are the statements that set forth the reasons or evidence, and the conclusion is the statement that the evidence is claimed to support or imply.'[3] For example:

P1 If you have broken the law then you will be punished.

P2 You have broken the law.

C Therefore you will be punished.

Often the premises come first in public presentations of arguments. Likewise, conclusions are commonly preceded by such 'indicator words' as 'therefore', 'hence', 'so' and 'it follows that'. And premises are identified by such indicators as 'given that', 'inasmuch as', 'for the reason that', 'since' and 'because'. But sometimes there are no indi-

cator words. And to decide whether or not they are dealing with a logical argument, the reader or listener must ask themselves such questions as: Is the writer or speaker trying to get me to believe something by giving me reasons for doing so? What are they trying to get me to believe? What reasons are they giving me?

Closely related to the concepts of *argument* and *statement* are those of *inference* and *proposition*.

> An inference in the technical sense of the term is the reasoning process expressed by an argument … A proposition is the meaning or information content of a statement.[4]

In making an inference we move from one or more pieces of information – A, B, C – to another piece of information – D. We see that A, B and C together 'point to' or 'imply' D. Logical implication is a relation between statements or ideas. A implies B where it really is the case that B follows from A. But it is humans or other intelligent agents who infer B from A, that is, who make the psychological step of drawing B from A, or seeing that, or how, B follows from A.

Our reasoning abilities are presumably grounded in deep structures of thought and perception that are products of our evolutionary and social history. Many animal species show evidence of elementary logical reasoning. But it is important that we all seek to sharpen up and improve our logical reasoning powers, so that our beliefs and actions are solidly based in facts and evidence, rather than prejudice, illusion and misconception.

Logic in law

It should also be clear that logic is relevant to law. Prior to their submission to parliament, proposals for particular bills are discussed and debated within government circles. Relevant evidence and arguments are considered. Interested outside bodies are often consulted. Law reform commissions and other bodies pursue research aiming to produce solid arguments in favour of particular changes to legislation and case law.

Proposals for Acts of parliament have to pass through lower and upper houses and become statutes only through the assent of both houses. Arguments for and against such proposed legislation are constructed, debated, modified and developed throughout the process.

Logical argument is equally relevant to ideas and practices of 'due process' in the day-to-day operation of the common law system. In civil proceedings plaintiffs have to assert their right to legal redress, and call for the issue of court orders against defendants. These arguments are set out in documents issued by court officers, and are met by counter-arguments (in similar documents) issued by defendants.

In criminal cases, the onus is upon the arresting (or summonsing) authority to provide the accused with justification of the charge in question. If the accused is charged with an indictable offence, a preliminary hearing may be held before a magistrate. The Crown produces its evidence in order to establish that it has a plausible ('prima facie') case against the accused. The magistrate is called upon to provide a logical assessment of the strength of the Crown case, to decide whether the evidence is sufficient, the argument strong enough, to warrant putting the accused to trial.

In civil cases, the onus is upon the plaintiff (or their representative) to persuade the court (judge or jury), through logical argument, that the facts they allege are true on a balance of probabilities, and together with relevant legal principles and precedents support and justify the issuing of a particular order of the court (typically requiring that they receive some sort of compensation from the defendant). Meanwhile, the defendant tries to refute such claims with logical arguments of their own. In criminal cases, the prosecution must prove the accused's guilt beyond reasonable doubt. A particularly strong logical argument is, or should be, required to establish such guilt.

The role of magistrate, judge or jury is that of assessing the strength of the competing arguments, for and against defendant or accused. This is sometimes said to be a largely 'passive' role. But it involves active construction of arguments to justify a particular verdict. Jurors will discuss and debate among themselves before arriving at some agreed structure of argument, though they are not required to provide public justification for their final conclusion. Judges are expected to provide written outlines of the processes of logical reasoning leading to and justifying their conclusions. In particular, they are expected to clearly identify established general rules or principles applied, and the justification for such application in the case in question. Starting law students learn to look for the 'ratio decidendi', that is, the 'rule of law expressly or impliedly treated by the judge as a necessary step in reaching his/her conclusion, having regard to the line of reasoning adopted by him/her'.[5]

Such written judgments, and transcripts of court proceedings, can, in turn, become the basis for appeals, involving logical criticisms of the arguments in question.

In all of these cases it would seem to be vitally important that good arguments prevail over bad. This depends upon the parties concerned being able to clearly and consistently distinguish good arguments from bad, to identify and criticise the bad and to develop and defend good arguments of their own.

If a lawyer sees their role as that of 'winning at all costs', and they have little in the way of valid or strong logical argument with which to do so, or believe that bad argument will be more effective than good in convincing judge or jury, then they could come to believe that they have a responsibility to utilise such bad reasoning. Such logically bad reasoning could be instrumentally or functionally good reasoning from their perspective, or that of their clients.

And if bad argument nonetheless serves to protect the innocent from wrongful punishment or ensure that the guilty get what they deserve, then it might appear difficult to say that it is not morally justified.

But in a reasonably run legal system, we should expect that bad argument, not recognised as such, is more likely to protect the guilty than the innocent. And lawyers' primary duty is supposed to be to the court, rather than to their clients. As Lord Denning MR said in *Rondel v Worsley*:

> The barrister must accept the brief and do all he honourably can on behalf of his client. I say honourably can because his duty is not only to his client. He has a duty to the court which is paramount. It is a mistake to suppose that he is the mouthpiece of his client to say what he wants or his tool to do as he directs. He is none of these things. He owes allegiance to a higher cause. It is the cause of truth and justice. He must not consciously misstate the facts. He must not knowingly conceal the truth. He must not unjustly make a charge of fraud, that is, without evidence to support it. He must produce all the relevant authorities, even those that are against him. He must see that his client discloses, if ordered, the relevant documents, even those that are fatal to his case. He must disregard the most specific instructions of his client if they conflict with his duty to the court.[6]

Particular defenders and prosecutors have achieved great fame and fortune effectively manipulating judges and juries through clever use of fallacious arguments (that is, bad arguments that look good), sometimes freeing the guilty and sending the innocent to punishment. On

the one hand, this is testimony to failure of relevant law societies to police the ethical codes of the profession. On the other, it demonstrates the failure of logical education of all parties concerned.

In an adversarial system the opposing counsel should be able to highlight, explain and refute such fallacious argument so as to prevent the judge or jury being taken in by it. And in the absence of such refutation the judges should intervene to disallow misleading argument.

Legislation or judge-made law introduced on the basis of poor justification and bad reasoning might, nonetheless, turn out to be good in its application. It may have good consequences for individuals or for society as a whole. But this would at best, largely be a matter of chance. At worst, such unjustified law-making would involve disastrous unforseen consequences or conspiracy by the legislators to favour particular groups at the expense of others without appearing to be doing so. Certainly, such unsupported or 'covert' law-making has nothing to do with genuine democratic choices of effective means to generally agreed ends.

Logic has undergone continuous development and evolution for more than 2000 years, so that, today, there are well-established principles for differentiating good and bad reasoning of all kinds. And yet contemporary Australian law texts seldom make any reference to such developments. Nor are law students in Australia generally expected or required to have any training in logic prior to graduation.

Where law textbooks do discuss logical reasoning their references are often outdated, confused or just plain wrong. We shall see some examples of this in what follows. Even the best regarded introductory treatments of logic in law – for example, 'The Reasoning of Lawyers' chapter in Waller's *Introduction to Law*[7] – are less than entirely clear in some areas.

Lawyers sometimes speak of there being a special kind of logic called 'legal logic', different from logic per se, or from logic as applied in other areas. It is true, as we will see, that there are many different kinds of logical arguments, with different criteria of strength or adequacy. And some sorts of arguments are more common or more significant than others in particular areas of human endeavour like science or maths or law. As we will see, analogical arguments are particularly important in law.

But there are no special, discipline-specific logics with their own unique rules. There is good logical reasoning and there is bad, that is, illogical, reasoning. The criteria for distinguishing good from bad ana-

logical reasoning are, in general terms, the same inside law as outside it. So if legal logic is not 'ordinary' good logic then it is simply bad reasoning masquerading as good. And, unfortunately, as we will see, there is far too much of such bad reasoning in law.

Some rules of evidence – for example, the general non-admissibility of prior convictions in a criminal trial – are specifically designed to reduce the likelihood of bad reasoning and fallacy in particular areas. But they are far from foolproof and are offset by what might be called the institutionalisation of bad reasoning and fallacy in other areas of the legal system.

It is all the more important, therefore, that starting law students do have a basic understanding of the nature of logical reasoning to guide them through this logical minefield. And a first crucial step here is to be able actually to distinguish logical arguments from the many other sorts of things people do with language and symbols.

Identifying arguments

Once we recognise that logical arguments involve the offering of reasons or evidence claiming to provide support of particular conclusions, we recognise that not all written or spoken communication involves logical argument. Such communication can involve alleged descriptions of people, things or situations, statements of belief or opinion, or explanations and value judgments, without any reasons or evidence offered in support of such descriptions, explanations or judgments.

The reader or the listener will frequently respond to such statements and judgments by considering possible evidence relevant to their truth or falsity. But for an argument to be present in the particular passage of discourse itself it must include 'a group of statements that can be analysed into premises and a conclusion, where an inferential claim is made that the conclusion, although controversial, should be accepted because of the evidence offered in the premises'.[8]

Upon first attempting to identify and analyse logical arguments – in textbooks, newspaper articles or legal judgments – it is easy to be misled by certain sorts of constructions that, at first, appear to be such logical arguments, but, in fact, are not. Here we consider two sorts of things frequently misidentified as logical arguments: conditional statements and explanations. We see how, while not themselves being logical arguments, they, nonetheless, play a central role in such logical argument.[9]

Conditional statements

So-called *conditional* or *hypothetical* statements, 'if X then Y', create particular problems insofar as they look like pairs of statements connected by indicator words. They do indeed include component statements, the one following the 'if' being called the 'antecedent', and the one following the 'then' the 'consequent'. But they are not logical arguments because there is no claim that either of these statements is true, or that one follows logically from the other.

As we have seen, in a logical argument, purportedly true (or probable) statements are presented as evidence or reasons to believe the truth (or probable truth) of another statement. 'In a conditional statement, there is no claim that either the antecedent or the consequent presents evidence … no assertion that either is true. Rather, there is only the assertion that if the antecedent is true, then so is the consequent.'[10]

But although conditional statements are not in themselves arguments, they play a very important part in logical reasoning and argument insofar as they express what are called *necessary and sufficient conditions*. A is said to be a sufficient condition for B whenever the occurrence of A is all that is needed for the occurrence of B, or the truth of A guarantees or necessitates the truth of B. B is said to be a necessary condition for A whenever A cannot occur without the occurrence of B, or A cannot be true unless B is true. In conditional statements antecedents identify sufficient conditions; for example:

- If a body is subjected to a net gravitational force (and in the absence of countervailing forces) then it will accelerate.
- If it doesn't accelerate then it hasn't been subjected to a gravitational force.

The force of gravity is sufficient to accelerate a body, but not necessary; some other force could achieve the same result (for example, an electrical force).

Consequents identify necessary conditions; for example:

- If you gained your LLB then you obtained a Pass mark (or higher) in your core law subjects.

Passing the core subjects is a necessary but not a sufficient condition for getting your degree. You also need to have passed your electives, paid your library fines and union dues, completed necessary paperwork, etc.

Another way of expressing a conditional statement is to identify the consequent (the necessary condition) by the phrase 'only if'.

- You will gain an LLB degree only if, etc.

And particular conditions can be both necessary and sufficient for other conditions:

- A solution is an acid if and only if litmus paper turns red when dipped in it.

Conditional statements often assert a causal connection between antecedent and consequent:

- If you heat a metal, then it expands.
- If the money supply is increased when all productive resources are fully utilised then inflation will result.

The term 'cause' can apply to both necessary and sufficient conditions and conditions that are both necessary and sufficient. Most often each of a number of necessary and jointly sufficient conditions are called *causal factors*.

When all other necessary conditions are present, a single necessary factor becomes sufficient to produce a particular result. Where other necessary conditions are relatively constant (the presence of oxygen, the availability of combustible materials), we are prone to identify a new development (the striking of a match), which together with such constant or background conditions is sufficient to bring about a particular effect or change (the fire) as 'the' cause. Where such a 'final' necessary condition involves some sort of human action, or something like human action, we are even more likely to identify it as 'the' cause.

Absences can be necessary conditions also. Just as the presence of oxygen and combustible materials are necessary conditions for fire, so are the absence of heavy rain or fire-retardant materials.

Lawyers are familiar with necessary conditions (and absences) treated as causes via the so-called 'but for' test. As Fleming says:

> The formula postulates that the defendant's fault is the cause of the plaintiff's harm if such harm would not have occurred without [but for] it. Thus a bather would not have drowned if a lifeguard had been present; the customer would not have fallen down the stairs if there had been a handrail; nor would he have suffered the brain haemorrhage if he had not received a blow on the forehead [the blow was a necessary condition of the haemorrhage]. Conversely, it is not a cause if the harm would have happened just the same, fault or no fault.[11]

But not all conditional statements express causal relationships. Another important role of conditional statements is to formulate *definitions*. For example, operational definitions involve specification of experimental procedures providing necessary and sufficient conditions for application of particular terms. We already encountered such a definition of the term 'acid' above. Here is another such definition:

- One substance is harder than another if and only if it scratches the other when the two are rubbed together.

Statements containing the expression 'if and only if' are called *biconditionals*. As Pine points out, such an expression is rare in everyday communication.

> However, such statements are found in logic, science, law, diplomacy, and any field where precise communication is very important. In 1979, the government of Iran told the US government, 'The US hostages will be freed *only if* the United States returns the Shah of Iran and his assets to the Iranian government.' It was important for US leaders to know the difference between this offer and 'The US hostages will be freed *if and only if* the Shah and his assets are returned.' The Iranian statement implied no guarantee that the hostages would be released, even if the Shah and his assets were returned. On the other hand, the use of *if and only if* would have been an implied guarantee.[12]

Many basic principles of law can be expressed in terms of a range of necessary conditions, jointly sufficient for a particular legal result: for example, forming a contract or establishing liability in negligence, or establishing various kinds of criminal liability.

- If it is established that (sufficient condition) (a) there is complete concordance between the parties as to the terms of an agreement; (b) the parties intend to be legally bound by their agreement; (c) the promises that constitute the agreement are supported by consideration; and (d) various other conditions are fulfilled, then (necessary condition) the courts will enforce the agreement as a contract.
- If a plaintiff is to make a successful claim in negligence then (necessary condition) they must be able to prove that the defendant owed them a duty of care, that the defendant breached that duty by failing to exercise the necessary level of care, and that the plaintiff suffered damage that was not too remote as a result of the defendant's breach.

- If the plaintiff is to establish that the defendant owed them a duty of care then (necessary condition) they need to establish foreseeability of harm and proximity between themselves and the defendant.
- If (sufficient condition) a person dishonestly appropriates property belonging to another with the intention of permanently depriving the other of it then (necessary condition) they have committed an act of theft.

These are definitions (or partial definitions) of particular legal terms (contract, negligence, theft). But they also express particular sorts of causal relations. In this case the agency of causation is the legal system itself, instituting 'appropriate' action if, and only if, it is established that particular social facts fulfil all of the necessary requirements for particular legal classifications.

Explanations

Another area that can create confusion is that of explanations. Like arguments, explanations involve two distinct components: in this case, the *explanans* and the *explanandum*. The former are the statements that do the explaining, while the latter describe the thing to be explained.[13]

Whereas a logical argument provides reasons for believing some otherwise doubtful claim, an *explanandum*, the thing to be explained, is typically some established fact. And rather than providing a reason for belief, the *explanans* aims to shed light upon the fact in question, typically by reference to the mechanism of its causation.

We noted earlier how causal relationships can be expressed through conditional statements. We can now see how such conditional statements can figure in explanations of particular observations or facts. We might, for example, want to know why railway lines have buckled.

Explanans

- If a metal is heated then it expands (all metals expand in proportion to heating).
- These metal railway lines were heated by extreme temperatures (and insufficiently large spaces were left between rails to allow for expansion beyond X centimetres).

Explanandum

- Therefore they expanded (beyond X centimetres, and buckled).

Much explanation involves such general principles or 'laws of nature' along with specific conditions of instantiation of such 'laws' (as illustrated above). And it is general principles of causation that typically provide the most useful explanations. But there are many different ideas about the appropriate nature and scope of explanation. As Burbidge notes:

> Causes produce results. And if we know what produced a state of affairs, we have some idea of why things are the way they are. Indeed, many people equate explanation with cause. If you give a cause and show how something was brought about, you have explained it … But there are other kinds of explanations as well … Another kind of explanation attempts to fit a number of different facts into a single, coherent picture. In Conan Doyle's classic detective story, 'The Speckled Band', we are told that Sherlock Holmes notices a number of details; the second half of a return ticket in the palm of his visitor's left glove; the left arm of her jacket spattered with fresh mud in no less than seven places. He puts these facts together to reach the explanation that Miss Stoner had started early to catch the train, and had a good drive in a dog cart along heavy roads before reaching the station. He has explained why a set of curious facts are just the way they are.[14]

Although explanations themselves are not arguments, we are certainly arguing when we provide reasons why others should believe that particular explanations are the correct ones. Lawyers in criminal and civil cases are often involved in arguing in favour of particular explanations of crucial facts. And expert witnesses are called upon to justify their (technical) explanations of particular relevant facts.

Another possible source of confusion here is that the same logical machinery involved in explanation, including appeals to natural laws and causal powers, can be used for purposes of prediction. There is always some doubt about what will happen in the future. And prediction is typically a form of argument, seeking to persuade the listener or reader that some specific thing will happen in the future.

This, of course, raises the question of what constitute good reasons for accepting particular explanations rather than others. Burbidge suggests that the fact that a particular explanation combines all the relevant facts into a single, integrated pattern provides good reason for accepting it. But we must be wary not to accept particular explanatory stories merely because they show how relevant facts can be coherently fitted together. As demonstrated in numerous legal proceedings, there will typically be many different stories that can be told around the same pieces of evidence (mud stains, tickets, etc). Lawyers are skilled

at weaving isolated facts into suitably coherent and appealing stories.

Here again, it is important to focus our attention upon the causal claims involved in the explanation and the means for supporting such claims. We have already considered an elementary understanding of causation in terms of necessary and sufficient conditions, and in chapters 5, 6 and 7 we will consider ways of discovering and establishing such conditions.

A deeper understanding of causation typically centres upon identifying and explaining the characteristic powers and tendencies possessed by particular sorts (or 'natural kinds') of things by virtue of their particular structures. We must then also consider the way in which particular powers are released or triggered to bring about particular sorts of results in particular sorts of situations. For example, by virtue of their cellular structure, seeds have the capacity to germinate and grow into plants, but such a power is released only in the presence of (such necessary conditions as) water, appropriate substrate and nutrients, including carbon dioxide and oxygen, appropriate temperature, and sunlight (or an appropriate substitute), and the absence of growth-restricting toxins or obstacles.

Metals have a power or propensity to expand when heated, because of the movement of free electrons inside the metal, behaving in similar fashion to the molecules of gas in a balloon. The attraction between the negatively charged electrons and positive ions in the metal hold the electrons trapped and create the effect of an impermeable but elastic sheath all over the surface of the piece of metal. The electrons are confined within this surface, with which they are continually colliding and recoiling. There is a resulting pressure on the surface of the metal. The average speed of the electrons increases as the temperature rises, thus causing the pressure to increase and the metal to expand.

Scientific experiment and observation provide logical support for general laws or principles relating to such powers and necessary conditions. The nature of the scientific reasoning involved in establishing such basic principles of causation is considered in some detail in later chapters.

In law, the principal causal agents are humans. We have already referred to the 'but for' test in the law of negligence. Here, a crucial issue is whether the defendant's negligent action (or inaction) really did cause the plaintiff's detriment. A defendant's action is judged to be negligent by reference to objective criteria of 'reasonable behaviour'. So the crucial causal question is typically whether a particular action really did produce a particular result. (Did lack of maintenance cause the gas storage tank to explode? Did the explosion kill the workers?)

In criminal law, a crucial issue is whether the accused's 'guilty mind' – or mens rea – has caused a particular criminal action – or actus reus. Here the law generally has to consider the subjective state of mind of the individual concerned: did they really intend to perform the action in question, and did their intention really cause them to perform that action? Legal issues of 'mind as cause' are considered in detail in chapter 18.

Arguments

Arguments can be categorised as *deductive* or *inductive* arguments.

> A deductive argument is an argument in which the premises are claimed to support the conclusion in such a way that if they are assumed true, it is impossible for the conclusion to be false. In such arguments the conclusion is claimed to follow necessarily from the premises. On the other hand, an inductive argument is an argument in which the premises are claimed to support the conclusion in such a way that if they are assumed true, then based on that assumption it is improbable that the conclusion is false (that is, it is probable that the conclusion is true). In these arguments the conclusion is claimed to follow only probably from the premises.[15]

It is often said that 'deductive reasoning moves from the general to the particular while inductive reasoning goes from the particular to the general'.[16] And it is certainly true – as we will see – that some important forms of deduction do move from the general to the particular and some important forms of induction in the opposite direction. But this is far from always being the case.

The distinction between deductive and inductive arguments lies in the strength of the argument's inferential claim, that is, in how strongly the conclusion is claimed to follow from the premises. Sometimes people are explicit in identifying the strength of the claims they are making. In this case, they use indicator words, prefacing the conclusion with 'necessarily' to indicate deduction or 'probably' to indicate induction.

Deduction

PI A square piece of land has sides that are 100 feet in length.

C Therefore, necessarily, it has an area of 10 000 square feet.

Induction

PI In a random sample of 1000 voters, 60% said they would vote Labour.

C Therefore, around 60% of all voters will probably vote Labour.

On other occasions, the interpreter has to rely upon the character or form of the argument and the actual strength of the inferential link between the premises and conclusion to decide the strength of the claim being made. The majority of mathematical arguments, like the one above involving an area, typically involve deductive claims. But, in the area of inferential statistics, where claims are made about the properties of populations (parameters) on the basis of the properties of samples (statistics), conclusions always follow with a certain probability, rather than with absolute necessity.

Every argument makes two basic claims: a factual claim, that the premises are true or probably true, and an inferential claim, that the conclusion follows, probably or necessarily, from the premises. If the premises fail to support the conclusion, it does not matter whether or not the premises are true: the argument is worthless. So logical-argument analysis starts with consideration of the inferential claim.[17]

Deduction

As noted above, arguments presented as *deductive* arguments are those in which it is claimed that the conclusion follows necessarily from the premises; in other words, it is claimed that it is impossible for the conclusion to be false if the premises are true. If the premises do in fact support the conclusion in this way the argument is said to be valid.

> A valid deductive argument is a deductive argument such that if the premises are assumed true, it is impossible for the conclusion to be false. An invalid deductive argument is a deductive argument such that if the premises are assumed true it is possible for the conclusion to be false.[18]

Arguments are often said to be valid by virtue of their (most specific)[19] form or structure. As Fisher says:

> The idea behind the notion of logical form is that one can distinguish between the 'structure' of an argument and its subject matter – or between its form and content. The content of an argument is what it is about (animals, atoms or whatever) and its form is expressed by means of those words which occur in reasoning about any subject whatever. If we call these the 'logical' words they include such examples as 'every', 'all', 'some', … 'no', 'if … then', 'implies', 'entails', 'follows from', 'because', 'so', 'therefore', 'and', 'but', 'or', 'not', 'is a' and many others. One can easily imagine a collection of arguments about

> very different subjects … which exhibit the same form when the words which are peculiar to each subject matter are replaced by neutral, schematic letters, A, B, C etc and one is left only with logical words.[20]

'An argument is said to be valid if it has a valid logical form, and a logical form is valid if there is no argument of that form which has true premises and a false conclusion.'[21] Certain structures of argument are always valid, whatever particular content we feed into them. Some basic valid forms of argument are as follows:

Modus ponens

P1 If A, then B. (If you dishonestly appropriated another's property then you are guilty of theft.)

P2 A. (You did.)

C Therefore, B (So you are.)

Modus tollens

P1 If A, then B. (If you are to be found guilty then you must have intended to permanently deprive the other of their property.)

P2 Not B. (You did not intend this.)

C Therefore, not A (So you are not guilty.)

Hypothetical syllogism

P1 If A, then B. (If you are found to have engaged in dishonest appropriation then you are guilty of theft.)

P2 If B, then C. If you are guilty of theft then you are liable for penal servitude.)

C Therefore, if A then C. (If you engaged in dishonest appropriation then you are liable for penal servitude.)

Notice that in each of these cases the first premise is a conditional or hypothetical statement of the kind considered earlier. For this reason they are called *conditional* or *hypothetical* arguments. But there are also other sorts of valid arguments:

Disjunctive syllogism

P1 Either A or B.

P2 Not A.

C Therefore, B.

A problem here is that we use 'or' to say two different things: at least one or the other and possibly both, or either one or the other but not

both. But in either case, this argument form remains valid.

By comparison, here is an example of a clearly invalid argument form:

P1 If A, then B. (If you are human then you are warm-blooded.)

P2 If C, then B. (If you a duck then you are warm-blooded.)

C Therefore if A then C. (So if you are human, then you are a duck.)

This conclusion clearly does not necessarily follow from these premises. It is quite possible for the premises to be true and the conclusion false, as in the example given here.[22]

There is no middle ground between validity and invalidity. Either an argument is valid or it is invalid. And there is only an indirect relationship between validity and truth. It is quite possible for an argument to have false premises and still be valid. Furthermore, a valid argument with false premises can have a true conclusion.

P1 All dogs are fish.

P2 All fish are mammals.

C Therefore, all dogs are mammals.

The point is that if all dogs were fish and all fish were mammals then it would necessarily follow that all dogs were mammals.

This highlights the fact that validity is a necessary but not a sufficient condition of a good deductive argument. Not only must the argument be valid, but its premises must be true (or well supported) also. Where the premises (used to support the conclusion) are true, and the argument is valid, it is called a *sound* argument. Otherwise it is unsound.

If the argument is sound then the validity of its structure transmits the truth of the premises to the conclusion: it too must be true. In the unsound argument above, the falsity of the premises means that the conclusion is unsupported. We are given no good reason to believe it, even though the argument is valid and the conclusion happens to be true.

Induction

In a valid deductive argument, if the premises are true, it is impossible for the conclusion to be false. If someone puts forward an argument as deductive, they are committed to the claim that this relationship holds:

that the conclusion follows necessarily from the premises. But this same relationship is not claimed to hold true for inductive arguments.

PI All swans we have observed are black. (This might be an observation by indigenous Australians prior to the arrival of the Europeans.)

C Therefore, all swans are black.

Here an inductive inference has been made from the characteristics of swans we have observed (a sample) to the characteristics of all swans, most of which we haven't seen. The reasoners realise that the premise of this argument could be true yet the conclusion turn out to be false. They state the premise as offering strong support for the conclusion; they do not claim that it necessitates the truth of the conclusion.

In a sense, valid deductive arguments say nothing more in the conclusion than is already said in the premises. But the conclusion of an inductive argument always goes beyond the information content of the premises; it always has a greater information content, and therefore opens itself to the possibility of being wrong, despite the truth of the premises.

If the premises do in fact provide such strong support for the conclusion, the argument is said to be (inductively) strong. A strong inductive argument is an inductive argument

> such that if the premises were true then based on that assumption it is probable that the conclusion is true. Conversely, a weak inductive argument is an inductive argument such that if the premises are assumed true, then based on that assumption it is not probable that the conclusion is true.[23]

Unlike validity, inductive strength is a matter of degree. And whereas a valid argument remains valid as we add further premises, strong inductive arguments do not necessarily remain strong with the addition of further information. If I add premises to the swan argument above to the effect that other competent witnesses report seeing numerous white and off-white swans, the earlier argument suddenly becomes much weaker.

An inductive argument that is strong and also has all true (or well-supported) premises is sometimes said to be a cogent argument.[24] Something the term 'strong' is used to cover this sort of case also.

Just as there are different forms of deductive arguments, so are there different forms of inductive arguments. Two of the most common are *analogical arguments* and *simple inductive generalisations*. The former

depend upon analogies or similarities between two or more things or situations. On the basis of such a similarity, a particular 'condition that affects the better known thing or situation is concluded to effect the similar, but lesser known thing or situation'.[25] The latter proceed from knowledge of a selected sample to some claim about the whole group of things of which the sample things are members.[26] Because all (or a percentage) of the sample things have a particular attribute, it is argued that all (or the same percentage) of the members of the group have the same attributes.

Simple arguments from analogy have the following structure:

P1 Entity A has attributes a, b, c, and z.

P2 Entity B has attributes a, b, c.

C Therefore, entity B probably has attribute z also.[27]

What matters in analogical argument is ensuring that the similarities between the things are appropriate to sustain the inference in question. Most important, the similarities must be relevant in the sense that a, b, c and z are somehow closely bound together in A, such that a, b and c cause z, or z causes a, b and c, or a, b, c and z are all produced by some common cause, or a, b, c and z are necessary structural features of a particular 'natural kind' of thing. So must we take account of the nature and degree of relevant disanalogy or dissimilarity between the entities concerned.

Burbidge provides a nice example of what happens without such a close connection:

P1 We can add, subtract, multiply and divide and we can write poetry as well.

P2 Calculators can add, subtract, multiply and divide.

C So calculators can [probably] write poetry.[28]

As Burbidge says:

> our ability to write poetry is more closely related to our imagination, our feelings, and our ability to sympathise with other people than to our ability to do arithmetical operations. And calculators are different from us in that they do not have imagination, feelings or sympathy.[29]

Analogical argument is itself very closely related to what we call creativity and imagination, once we include references to dissimilarities as well as similarities.

P1 A's are like B's in attributes a, b, and c, but different in attributes d, e and f.

P2 B's also have attributes g.

C So possibly A's have something like attribute g, suitably modified to take account of the differences, d, e and f.

P1 Mountains are like molehills in shape, location and material content, but radically different in size (that is, much bigger).

P2 Molehills are made by moles.

C So possibly mountains are made by something like moles, but radically different in size (that is, much bigger).

Through the simple (mechanical) process of juxtaposing similarities and dissimilarities we have generated a strange new idea, seemingly out of nowhere: that of a giant, mountain-generating mole. It is easy to apply this process in other cases, though with no guarantee of useful results.

We have already encountered an example of simple inductive generalisation in the earlier argument about swans. Such simple generalisation becomes increasingly organised and reliable in the science of statistics. The crucial issue here is ensuring that the sample is representative of the population. Samples that are not representative are said to be biased. And inferences from biased samples are not at all reliable. As Hurley notes:

> depending on what the population consists of, whether machine parts or human beings, different considerations enter into determining whether the sample is biased. These considerations include (1) whether the sample is randomly selected, (2) the size of the sample, and (3) psychological factors.[30]

Lawyers in the modern world need to know a little bit about statistics in order to be able to make sense of scientific research and expert testimony. We will return to this point in later chapters.

Hurley also points out:

> Analogical arguments are closely related to generalisations. In a generalisation, the arguer begins with one or more instances and proceeds to draw a conclusion about all the members of a class. The arguer may then apply the generalisation to one or more members of this class that were not noted earlier. The first stage of such an inference is inductive, and the second stage is deductive. For example, a man thinking of reading a Stephen King novel might argue that, because the last three

> novels by King were thrilling, all King novels are thrilling [induction]. Applying this generalisation to the latest novel, he might then conclude that it too is thrilling [deduction]. In an argument from analogy … the arguer proceeds directly from one or more individual instances to a conclusion about another individual instance without appealing to any intermediate generalisation. Such an argument is purely inductive. Thus the arguer might conclude directly from reading three earlier King novels that the latest one is thrilling.[31]

Standard of proof

In criminal law, the prosecution is supposed to prove their case beyond a reasonable doubt in order to secure a conviction. This does not mean they must prove their case with deductive (mathematical) certainty. Even if the relevant principle of law is clear (if A, then B), the facts (A) have still to be established. And this will always involve some form of inductive, and hence probabilistic, argument. Even if the criminal is (apparently) clearly identified on the surveillance video, there is still some doubt: it could be someone wearing a mask designed to look like someone else.

This raises the question of what degree of probability – or likelihood – counts as 'beyond a reasonable doubt'. In some scientific applications 95% is regarded as good enough to 'refute the null hypothesis', equivalent to refuting the presumption of innocence. In others, 99% probability is required.

In civil law, the required standard of proof is considerably weaker. The plaintiff succeeds in establishing the defendant's liability if their case is stronger 'on the balance of probabilities' (or balance 'of the evidence'). This presumably means that even a tiny advantage in terms of probability wins the decision. This has its own problems, especially considering how much could hang on the decision. Fifty-one per cent likelihood hardly seems much better than 49%. And what if neither case looks particularly convincing?

Extended arguments

All of the arguments considered so far have been *simple arguments*, which is to say, single sets of premises offered in support of single conclusions. In day-to-day social life, and certainly in legal proceedings, real arguments are often *complex* or *extended arguments*. Extended arguments are interconnected systems of simple arguments where the conclusions of one or more such arguments become inputs – as prem-

ises – into other such simple arguments. Such conclusions, which are also premises, are called *intermediate conclusions*, that is, stages on the way to a *final conclusion*. In this way, complex chains and trees of argument are developed, with different branches or tributaries feeding into or converging upon a single final conclusion.

As we will see, legal procedures typically involve such interconnected chains of argument. Factual statements established as true conclusions of inductive arguments become premises in legal arguments about the correct legal decisions and responses.

It is often difficult to sort out the structure of complex and lengthy extended arguments. It is useful to number all of the relevant statements and then draw diagrams to show the logical relations between the relevant numbers. Premises go at the top, either on their own or linked to other premises (depending upon whether they are taken to imply conclusions alone or in combination), with downward pointing arrows to their conclusions. Different sorts of arrows can be used to distinguish deductive from inductive steps.

The situation is further complicated because not all logical arguments – outside the pages of logic textbooks – are presented in fully explicit form. Intended conclusions and/or one or more of the premises might not actually be stated. Such arguments, with components only implicitly suggested, are called *enthymemes*. Typically conclusions are left out because they are presumed to be obvious from the context. Premises are left out because the arguer assumes their audience is already aware of, and accepts, the facts in question.

As a general rule, in reconstructing and assessing arguments, we can apply what has been called a *principle of charity*, giving the arguer the benefit of the doubt in supplying whatever missing statements are needed to make the argument as strong as possible. Consider the following argument:

> (1) If God can create a stone too heavy for Him to lift, there is something He cannot do; (2) and if He cannot create a stone too heavy for Him to lift, there is something He cannot create. (3) If there is something God cannot do then He is not omnipotent; (4) and if there is something He cannot create He is not omnipotent. (5) Therefore God is not omnipotent.

From statements (1) and (3) it follows that if God can create a stone too heavy for Him to lift, then He is not omnipotent. From statements (2) and (4) it follows that that if He cannot create such a stone He is

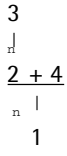
not omnipotent. These are unstated intermediate conclusions. And it therefore follows (as another unstated intermediate conclusion) that if He can or cannot create such a stone He is not omnipotent. An unstated – but obviously true – premise here is that it is true that He either can or cannot do so: there is no third possibility. So the stated conclusion does indeed follow, by *modus ponens*.

In analysing any complex extended argument it is often helpful to start by looking for the beginning and the end of the argument. What is being assumed to be known or accepted here (where does the argument start)? And what is the final conclusion (where does it end)? Then we look for intermediate conclusions and the evidence offered to support them. Here is a simpler example:

> (1) When parents become old and destitute, the obligation of caring for them should be imposed on their children. (2) Clearly, children owe a debt to their parents. (3) Their parents brought them into the world and cared for them when they were unable to care for themselves. (4) This debt could be appropriately discharged by having grown children care for their parents.[32]

In this case, we see that the final conclusion is stated first: the arguer is trying to convince their audience that children should be the ones to take responsibility for care of their aged parents. Then we ask, are there any statements presented merely as matters of fact without further justification or reference to other parts of the argument? And we see that there is one such statement, statement (3), to the effect that parents cared for their children when such children couldn't care for themselves. This is therefore a premise, offered as a ground or reason for believing statement (2), that such children, therefore, owe a debt to their parents. Statement (2) is therefore an intermediate conclusion on the way to the final conclusion.

Statement (4) is also presented as fact, and is also only a premise, but in this case it makes reference to the debt that has (supposedly) been established to exist in the sub-argument 3 – 2. This tells us that it is 4 and 2 together that are presented as justification for the final conclusion, 1. The overall structure of the argument is this:



Here is another (mathematical) example:

> (1) The whole number series is infinite. (2) If it weren't infinite, then there would be a last or highest number. (3) But by the laws of arithmetic, you can perform the operation of addition on any arbitrarily large number; call it n to obtain n + 1. (4) n + 1 always exceeds n. (5) So there is no last or highest number (6) and the series of integers [whole numbers] is infinite.[33]

Here, the final conclusion appears at both the beginning and the end. Statements (3) and (4) stand together to support statement (5), and statements (2) and (5) together support the final conclusion, statement (1).

$$\begin{array}{l} \underline{3 + 4} \\ \quad {\scriptstyle n} \quad | \\ \underline{2 + 5} \\ {\scriptstyle n} \ | \\ \quad 1 \end{array}$$

Here we see that the second sub-argument, 2 + 5 = 1, has the form of *modus tollens*:

P1 If not I, then H.

P2 Not H.

C Therefore, not not I. [That is, I]

Here is a slightly more complicated example, taken from a letter to a daily newspaper. Individual statements have been numbered to facilitate logical analysis.

> [1] In rejecting new powers for police in the fight against drugs civil libertarians apparently believe that it's better to have 100 drug pushers on the street than a single innocent person in jail. [2] They have difficulties with police conducting medical examinations of those suspected of hiding drugs inside their person, and drug dogs deployed against the public in random searches, [3] but this is exactly what Australian Customs Service does every day at Sydney airport. [4] The only solution the libertarians offer is education, [5] but no matter how well people are educated they can still be tempted into unlawful and anti-social practices. [6] The temptation to use drugs is only possible if drugs are available. [7] The new police powers are clearly aimed at suppliers and their supply lines. [8] The law-abiding public want the drug problem cleared up. [9] Posturing, barrow-pushing civil libertarians with no solutions are part of the problem and should get out of the way.[34]

In this case readers are encouraged to work out their own structure diagram.

Discussion topics

1. What is logical reasoning? Explain these concepts: arguments, premises, conclusions. What is the role of logical reasoning in law?
2. Discuss some major problems in recognising arguments. Explain conditional statements, explanations and arguments.
3. Explain deduction and induction. Include reference to validity, truth, soundness, inductive strength, *modus ponens*, analogical reasoning and generalisation.
4. What are extended arguments? Include examples. Refer to textbooks, newspaper editorials and leader articles and judgments.
5. What is analogical reasoning? Why is it important?

For preliminary consideration of legal reasoning see Lord Keith's judgment in *Hill v Chief Constable of West Yorkshire* (1989) 1 AC 53 (House of Lords) as reproduced in J. Swanton, B. McDonald and R. Anderson, *Cases on Torts*, Federation Press, Sydney 1992, pp 135–8. What kinds of argument are involved here? Are they strong?

See also Panelli J's judgment in *Moore v Regents of the University of California* (1990) 793 P 2d 479 (Cal Sup Ct), along with Mosk J's dissent, as reproduced in I. Kennedy and A. Grubb, *Medical Law: Text and Materials*, 2nd edition, Butterworths, London, 1994, pp 1112–24. What are the key arguments? Who is correct?

Chapter 2

LEGAL REASONING

Law reform

All the sorts of arguments considered so far, deductive and inductive, occur in legal contexts. But different sorts of arguments predominate depending upon whether we consider the creation of law or its application.

The traditional view is that laws exist to fulfil particular social functions. So at the law-making stage the central questions are: What issues that we might reasonably expect law to be able to address are most urgently in need of addressing? What sorts of laws are capable of effectively addressing such issues?

Thus, in government departments, in parliaments and in organisations involved in lobbying and law reform, logical arguments are developed to demonstrate how and why particular social issues or problems need to be given priority. And other arguments are developed to demonstrate how particular changes or additions to the body of established law can be expected to effectively address such issues or problems, without creating greater problems of their own. Often, such arguments give rise to counter-arguments and counter-proposals of various kinds.

In broad terms, issues of social needs and priorities are moral issues. Why are some rights or interests in more urgent need of protection than others? Why are some duties in more urgent need of enforcement? Why and how should we aim to achieve the greatest good for the greatest number? Why and how best should we protect those rights and interests are threatened?

Clearly, there are also economic issues involved here. Particular rights and interests cannot be prioritised, protected or enforced without the material wherewithal to do so. There will always be questions of how such reforms are financed and who pays. We will look in greater detail at the nature of such moral and economic issues and arguments in later chapters.

Issues of the practical efficacy of proposed legislation in effectively addressing such social problems are basically scientific issues. They depend upon an understanding of the causes of the problems in question, and of the likely consequences of particular sorts of legal interventions. We will look in greater detail at these sorts of issues and arguments also in later chapters.

A stark example of the interaction and interdependence of ethical and scientific issues in possible law reform is provided by ongoing debate about the continued application of the death penalty for first degree murder in 38 of the 50 US states (along with the military and the US federal jurisdiction). On the one side are straightforwardly factual issues of determining the causal efficacy of capital punishment in reducing homicide rates (through effective deterrence) or increasing these rates (through legitimating violence or other considerations). Here too are factual issues of discriminatory application of capital punishment (through racist and class-ist attitudes on the part of police, lawyers, judges and juries and inadequate access to legal resources on the part of black or poor Americans) and numbers of innocent people wrongfully convicted and executed.[1]

On the other side are ethical issues, with some ethical approaches finding the death penalty intrinsically unacceptable, others finding it intrinsically appropriate and just retribution for first-degree murder, and others again – specifically utilitarians – seeking to assess it in terms of the overall balance of pain and happiness (or social welfare) achieved through retention or abolition.

The significance of factual data for the utilitarian position is immediately apparent. Utilitarians cannot make judgments about the quantitative balance of overall pleasure and pain – or social benefit and detriment – without reliable evidence of the quantitative consequences of capital punishment. They need to know about the extent of effective reduction (or increase) in homicide rates, of inequality of application of the death penalty and miscarriages of justice.

We might not think that factual issues would be so relevant for the other two sorts of ethical perspective. But, here again, factual data can

be of crucial significance. Those who find the death penalty intrinsically unacceptable have a great interest in determining the real causes of homicide, so as to be able to take effective action – other than capital punishment – to reduce it. And even those who believe in retribution need to seriously consider issues of radical inequality of application of capital punishment (with a 40-times greater chance of a death penalty being sought against someone convicted of killing a white person than a black person) and of executions of the innocent. It is, after all, quite possible that they put a significant value upon fairness and protection of the innocent as well – possibly a higher value than they put on retribution.

In Australia, the death penalty is – supposedly – no longer an issue. But precisely similar considerations apply to incarceration within the prison system as the most serious criminal penalty, especially given the high levels of homicide and fatal disease in Australia's prisons.

It is all too easy in our current system for serious moral deliberations to be subverted by powerful minorities, concerned only with their own interests in expanded wealth and power. In a capitalist society certain groups – particularly business leaders – are in a position to effectively influence law-making in their own favour – and at the expense of other groups – through their organised political lobbying,[2] their contributions to political parties, their threats and promises in relation to job creation, and their control of the mass media. But in a supposedly open and democratic society such powerful minorities (or their political representatives) will typically try to disguise such pursuit of narrow self-interest with a veneer of moral concern for the 'general will', for the good of the majority, or for the good of oppressed and suffering minorities. We examine these issues further in the second and third sections of the book, dealing with Law, Ethics and Social Theory and Contemporary Social Problems.

Nor do the deliberations of law-makers necessarily centre upon real scientific understanding of the issues in practice. Cynical politicians may push through repressive laws in order to win easy votes, knowing full-well that such laws will not achieve their stated aims and/or will create further serious social problems. Or else the science is covert, with laws enacted to supposedly address matters of general concern, when in fact they have been carefully constructed to serve the narrow interests of powerful minorities.

All this applies first and foremost to legislation, rather than to judge-made law, although courts can be subject to similar economic

and political pressures, at least indirectly. In practice, judges make law (cannot help making law) in the course of application of existing law. And it is to such application that we now turn.

Arguments in court

As noted earlier, general legal principles (expressed in statutes or judgments of common law), in the form of conditional statements, play a central role in assigning particular sorts of facts to particular legal concepts or categories, and identifying appropriate (accepted) legal responses to the facts in question when so assigned. The prosecution or the plaintiff is claiming that certain things happened and that those things have a certain legal significance, calling for a particular legal response. The accused or the defendant will be contesting these claims.

Much legal discussion and argument therefore revolves around issues of what actually happened in the circumstances identified by prosecution or plaintiff, and involves inductive-logical arguments supporting or refuting particular factual claims. This will include arguments relating to eyewitness testimony and expert testimony.

If an eyewitness claims to have seen certain things, then evidence that they are a truthful and competent witness, that they have no reason to lie in this case, that they were indeed in a position to witness the events in question and were unlikely to have misunderstood the significance of what they saw, counts in favour of the likelihood that the events in question did indeed take place. This is, of course, inductive reasoning, since we can never be absolutely certain of what really happened.

Experts will typically be called in to testify about the sorts of events likely to have produced particular items of evidence including such material evidence as bloodstains, tracks, and DNA, as well as the nature and causation of injuries or illnesses (including mental illnesses) on the part of accused persons, their victims and plaintiffs. They will typically be called upon to explain the evidence in question, by reference to some mechanism of causation. They might, for example, explain the behaviour of the accused in terms of a particular psychopathology, or the disease symptoms of the plaintiff in terms of exposure to a particular toxin or pathogen.

Here, the mere fact that such witnesses are accredited experts in a relevant area carries some weight in supporting their assertions.

P1 Most accredited experts can be relied upon most of the time to form true or probable opinions in the areas of their special knowledge and expertise.

P2 This is an accredited expert.

P3 They are testifying in an area where they have special knowledge.

C So, probably, we can assume that what they say is true.

But experts are (or should also be) called upon to justify their explanations by appropriate logical argument. Typically this means justifying the general principles involved in such explanations by reference to the scientific evidence in support of such general principles.
They might refer to clinical, epidemiological or laboratory evidence providing support for their causal claims. This might include experiments in which a certain percentage of rats or dogs subjected to a particular toxin in certain amounts for certain periods developed some form of cancer, and so-called prospective studies of particular human populations exposed to high levels of the toxin in question (and compared with control groups, not so exposed).

There is also legal argument about whether – agreed – facts are really the kinds of facts that figure as necessary (or sufficient) conditions in general principles of law (statutes, by-laws or principles of common law).

The doctrine of stare decisis, to the effect that cases involving the same essential or material facts must be decided in the same way, ensures that analogical argument plays a central role in the common law. For no two cases are exactly alike, and all cases are alike in some way or another. So, even with agreement as to the facts of the present case, there are always questions of which previous cases are relevantly similar to the present case and therefore define the appropriate response in the present case also.

In an adversarial context, each side will typically seek to assimilate the present case to past cases with an outcome that would be favourable to them if applied in the present case. This means competing analogical arguments. Muragason and McNamara describe the circumstances in *R v Easom* [1971] 2 QB 315:

> D had taken V's handbag while sitting in a cinema and had examined it to see if there was anything worth stealing. D decided there was nothing worth stealing and replaced the handbag. D was arrested and charged with theft. But the Court of Appeal stated that in every case of theft the appropriation must be accompanied with an intention of

> permanently depriving the owner of the property; 'conditional appropriation will not suffice'. Lord Edmund-Davies LJ, delivering the judgment of the court, held (at 319), 'if the appropriator has it in mind merely to deprive the owner of such of his property as, on examination, proves worth taking and then, finding that the booty is valueless to the appropriator, leaves it ready to hand to be re-possessed by the owner, the appropriator has not stolen'.[3]

In the later case of *Sharp v McCormick* (1986) VR 869, D was caught by the police in possession of a motor car coil that he admitted having taken from his employer dishonestly and without permission. D claimed that he was intending to fit it to his car. If it worked D intended to keep it. However, if it did not work he planned to return it. He argued that he had only a conditional intention and that the decision in *Easom* governed the case.

P1 This case is like *Easom* in that there was a merely conditional appropriation of the property of another.

P2 In *Easom* the accused was found not guilty of theft (it was not a case of theft).

C So the accused should be found not guilty in this case as well (this also is not a case of theft).

The prosecutor, by contrast, sought to assimilate the facts of this case to those of others where the actions in question had been deemed to constitute theft. And this was supported by the decision of the Supreme Court of Victoria. The judges distinguished the case from *Easom* by reference to D's intentions at the time of appropriation, arguing that at that time he was 'clearly treating the coil as his own to dispose of as he saw fit', unlike the accused in Easom, who did not so treat the contents of the handbag. 'He may or may not have' returned the coil if it didn't fit, whereas the accused in *Easom* had returned the contents of the handbag.

In other words, they appealed to (what they claimed to be) a significant disanalogy between the two cases: sufficiently significant to support a guilty verdict in this case. They argued that the 'essence' or true nature of theft is such that *Easom* falls outside of the concept while *Sharp* falls inside. And all of this can be seen as modifying, developing or clarifying the definition of theft; the necessary conditions for determining what particular sort of material facts count as theft.

In many cases, issues revolve around questions of *statutory interpretation*. As Hall notes, 'It is suggested that courts are required to rule

upon the meaning of legislation in approximately 50% of all cases.'[4]

One side might argue that the word 'false' in a particular Act dealing with false statements about the importation of goods means any statement that is found not to have been true. The other side might argue that the Act in fact refers only to the making of purposefully untrue statements by the person concerned, not to simple errors of fact.

The fundamental issue here is the ambiguity of the words used in legislation (the precise scope of their reference), and the different possible techniques available to resolve such ambiguity.

> While the starting point for the court is still the ordinary meaning of the words used in legislation, the courts are more willing [than once they were] to consider the outcome that results from a particular interpretation in deciding whether ambiguity exists.[5]

And they increasingly look to the purpose underlying legislation in order to resolve such ambiguity.

Section 15AB of the *Acts Interpretation Act 1901* (Cth) allows consideration of such 'extrinsic' evidence as records of parliamentary debates and speeches, explanatory memoranda, executive documents, reports by parliamentary committees and commissions, and international agreements in order to clarify the purposes underlying the legislation where a provision is ambiguous or obscure, or where the 'ordinary meaning' leads to a result that is manifestly absurd or unreasonable.

Various common law guidelines and presumptions are available for resolving problems of statutory interpretation, but as Hall says, these are not now seen as overriding the purposive approach, based upon clarifying the intentions or purposes of the legislators.

The concluding statement of a prosecutor or counsel for the plaintiff is likely to take the form of a deductive, *modus ponens*-type argument: Given that X and Y are the kinds of facts referred to in a particular principle of law, as necessary and jointly sufficient conditions for a particular legal state of affairs, and for a particular legal response, and given that X and Y have been found to pertain in this case, then it necessarily follows that judge or jury should draw the appropriate legal conclusion and take the appropriate action.

> PI If A and B (particular types of fact are found to pertain), then C (a certain type of legal situation is established to pertain) and D (certain types of legal response must follow).

P2 A and B (the sorts of facts in question have been established to pertain in this case).

C Therefore C and D (the relevant legal response must follow).

The first premise may be true as a matter of law, and the argument might be valid, but the crucial question is whether it is sound: in other words, whether the second premise is also true (or well supported). This is where the jury is called upon to make a judgment of fact.

The concluding statement of counsel for the accused or the defendant is likely to take the form of a deductive, *modus tollens*-type argument: Given that facts of type X are necessary conditions for guilt or liability in this sort of case, and that it has not been established – beyond reasonable doubt or on the balance of probabilities – that a type X fact is involved in this case, then it necessarily follows that judge and jury should reject the claim of the other side and dismiss their case.

P1 If C and D (if particular legal consequences are to follow), then A (a particular type of fact needs to be proved to pertain).

P2 Not A (this type of fact has not been proved to pertain in this case).

C Therefore (not C and D).

Judgments

Finally, there are written judgments delivered by presiding judges. And it is here, particularly in appeals to higher courts, that we most clearly see the judges in the process of making law, through inductive generalisation and analogy.

Such judgments typically start with identification of what are taken to be the relevant facts, followed by identification of what is taken to be the relevant category of law. The issues in dispute within that area are identified, and relevant cases and/or legislation cited. Some relevant general principles are derived – by inductive generalisation from past cases, or by direct reference to statements of such common law principles in previous cases or to legislation. Such general principles are sometimes analysed – and justified – in terms of broadly ethical or social considerations. Qualifications to such general principles are then considered, including issues of 'policy' leading to the formulation – and application – of a more specific rule, dealing with the application of the general principle to the particular sort of case at issue.

A straightforward example of identification and application of an existing rule is provided by *Hirachand Punamchand v Temple* (1911)

2 KB 330. The rule in question does not allow a creditor to sue a debtor for the remainder of a debt if that creditor has agreed to accept a lesser amount from some third party. In this case, moneylenders Hirachand sued Temple, to whom they had lent a sum of money after cashing a cheque from the debtor's father who had stipulated that said cheque was in full settlement of the debt. In the Court of Appeal, Fletcher Moulton LJ argued as follows:

> If a third person steps in and gives a consideration for the discharge of the debtor, it does not matter whether he does it in meal or in malt, or what proportion the amount given bears to the amount of the debt. Here the money was paid by a third person, and I have no doubt that upon the acceptance of the money by the plaintiffs with full knowledge of the terms on which it was offered, the debt was absolutely extinguished (at 340).[6]

It is easy to see the basic *modus ponens* form of argument employed here:

P1 If a third party offers some amount in full discharge of a debt and the creditor accepts this amount then the debt is absolutely extinguished.

P2 In this case both of these conditions were fulfilled (in cashing the cheque the creditor 'accepted' the money in question as discharge of the debt).

C Therefore in this case the debt was absolutely extinguished.

Classic examples of formulation of new principles through inductive generalisation are *Heaven v Pender* (1883) 11 QBD 503 and *Fletcher v Rylands* (1866) LR Exch 265. In the former case, 'the plaintiff, a painter, had been injured when some defective scaffolding, supplied by the defendant to the plaintiff's employer, collapsed'.[7] Brett MR reasoned that there were many separate classes of cases in which persons owe a duty of care not to injure others. Specifically he mentioned four such classes of cases: 'when two drivers [of carriages] or two ships are approaching one another', the case of a railway company carrying a passenger, and cases 'with regard to the condition in which an owner or occupier leaves his house or property'.[8] From such cases he drew out the following general principle:

> Whenever one person is by circumstances placed in such a position with regard to another that every one of ordinary sense who did think would at once recognise that if he did not use ordinary care and skill in his own conduct with regard to those circumstances he would cause danger of injury to the person or property of another, a duty arises to use ordinary care and skill to avoid such danger.[9]

In the latter case:

> Rylands and his partner had built a water-storage dam, which happened to be over some old, filled mine workings. Water from the dam broke through the old shafts and flowed through Fletcher's mine, still lower down. Blackburn J on behalf of the whole court reasoned that as there were separate rules imposing strict liability for cattle trespass, the escape of filth from privies, and the escape of noxious gases and smells, there was a general rule that whoever brought onto his or her land something which by its nature was liable to escape and cause damage was strictly liable when it did so. He then applied this rule to the escape of water in this case.[10]

The decision in *Fletcher v Rylands* was confirmed on appeal by the House of Lords, but Brett MR's argument was rejected by the other trial judges who saw his general principle as in conflict with an already established principle that attributed a duty of care to a seller of a good only to the extent that such a duty had been established through a contract of sale with the purchaser of the good in question.

Privity of contract provided that someone not a party to a contract could not sue for breach of that contract, even if they had suffered harm through the defective character of the goods in question.

The only exception (following the 1851 decision of *Longmeid v Holliday* 6 Ex 761; 155 ER 752) was articles said to be 'inherently dangerous'. But this situation changed profoundly with Lord Atkin's famous speech in *Donoghue v Stevenson* (1932) AC 562, which defended Brett MR's reasoning as basically correct, once his 'too wide' principle was narrowed down to produce the famous 'neighbour principle' as the 'general conception of relations giving rise to a duty of care, of which the particular cases found in the books are but instances' and the 'common element' upon which all liability in negligence was 'based':

> The rule that you are to love your neighbour becomes in law, you must not injure your neighbour, and the lawyers' question, Who is my neighbour? receives a restricted reply. You must take reasonable care to avoid acts or omissions which you can reasonably foresee would be likely to injure your neighbour. Who, then, in law is my neighbour? The answer seems to be – persons who are so closely and directly affected by my act that I ought reasonably to have them in contemplation as being so affected when I am directing my mind to the acts or omissions which are called in question.[11]

From this general principle, along with the facts of the case, is then derived a more specific rule determining the outcome in this case:

> By Scots and English law alike a manufacturer of products, which he sells in such a form as to show that he intends them to reach the ultimate destination in the form in which they left him with no reasonable possibility of intermediate examination, and with the knowledge that the absence of reasonable care in the preparation or putting up of the products will result in an injury to the consumer's life or property, owes a duty to the consumer to take that reasonable care.[12]

From this point, and along with the facts of the case taken to be established, the decision becomes a straightforward matter of deduction.

It is also important to note, with MacAdam and Pyke, that 'in the cases where judges announce a new principle or a broad rule' including all the cases just considered, 'they characteristically appeal both to a number of precedents and to moral notions or popular thought, as evidence of the existence of the principle'.[13]

In *Fletcher v Rylands*, Blackburn J speaks of his general rule as 'reasonable and just'. In *Heaven v Pender*, Brett MR appeals 'not only to 'the logic of inductive reasoning' but to 'the universally recognised rules of right and wrong'.[14] And in *Donoghue v Stevenson*, Lord Atkin appeals not only to 'logic' but also to 'a general public sentiment of wrongdoing' to 'sound common sense' and to the biblical injunction to 'love your neighbour as yourself' (Leviticus 19; 18). The moral principles themselves are not justified but are rather seen as offering independent support for the new legal principles.

Logic, law and history

There is a long tradition of thought arguing that the common law is grounded in, and logically derived from, a handful of general principles, handed down from the remote past. According to the 'declaratory' doctrine of the common law, judges do not make law, but merely transmit, interpret and apply it. As Blackstone said, they are 'the depositories of the laws, the living oracles, who must decide all cases in doubt'.

Such interpretation and application involves the application of deductive and inductive reasoning. But it does not involve any sort of principled or creative response to changing historical circumstances in the light of personal moral or political commitments.

Yet the extension of established precedent and principle through analogy depends, as we have seen, on principles of relevance that themselves change and develop through time. With changing social and political consciousness, new principles of relevance come to the

fore. As we have seen, inductive generalisations like the neighbour principle always go beyond specific cases – and groups of cases; differently selected cases or the same cases can sustain a range of different generalisations.

The historical forces underlying the logical development of the law of negligence, just considered, are complex and cannot be properly explored at this stage. There is general agreement that the increasing pace of the industrial revolution throughout the 19th century played a central role, increasing the range and frequency of serious accidents involving strangers, with victims (increasingly isolated from traditional social support systems) turning to the legal system for redress. The second half of the century also saw an increasing ideological shift away from the extreme libertarian individualism of the early industrial revolution to more social-liberal thinking, including utilitarian ideas of pursuit of general social welfare as an intrinsic good, legislation dealing with health and safety at work, legalisation of trade union activity, and the extension of the franchise, with universal male suffrage in Britain by 1867.

These changes were arguably bound up with substantial political movements affecting broad layers of the population – Chartism, the growth of trade unionism, socialism. Commentators have also pointed to the influence upon jurisprudential thinking of the rapid pace of scientific advance in the later 19th and early 20th centuries, supposedly achieved through application of inductive generalisation.[15]

All of these processes brought pressure to bear for the application of inductive reasoning to 'a residual category-less number of non-contractual, non-criminal wrongs encompassing the torts of trespass and case' in order to produce 'a comprehensive theory of universal applicability' recognising the rights of individuals to the due care of their fellows over and above the purchase of such care through contractual arrangements.[16]

As we have seen, Brett MR sought to achieve this goal in 1883, but failed to win the support of fellow judges. It is likely that at this stage, in Britain, such a principle was seen as too much of a threat to 'national economic development', threatening the profits of industrial entrepreneurs through actions by injured workers and consumers. Indeed, from the mid-1800s the law of negligence seems to have been an effective means through which the legal profession ensured the active subsidisation of industry by 'shifting the costs of

entrepreneurial activity from industrialists and on to workers, consumers and bystanders'[17], through such common law doctrines as contributory negligence, voluntary assumption of risk and common employment.

However, in the course of the later 19th century, all of these restrictive principles were significantly relaxed or abolished. And by 1932, the time of *Donoghue v Stevenson*, much more social-liberal ideas were in the ascendant, following further massive extension and consolidation of the power of the industrial middle class, and the Great Depression, indicating the clear need for greater regulation and accountability on the part of industrial capitalists.

We will return to consider some of these issues and ideas in greater depth in later chapters. For present purposes, the important point is that they in no way detract from the importance of logic in the common law, or of students' understanding of logic in understanding the historical development of the common law.

Logic is the means through which legal practitioners respond to the historical forces constantly impinging upon them. And we must be able to follow the threads of such logical reasoning to understand precisely how this is happening, to see where their true priorities lie and whose interests they are really trying to serve. In particular, we must be able to see how bad legal reasoning frequently serves to mask a hidden political and social agenda, a covert system of political and social priorities and reasoning based upon such priorities, beneath the surface of the public justifications and rationalisations.

Discussion topics

1. What is the role of logical reasoning in law reform?
   ...in court proceedings?
   ...in judgments?
2. What is the relationship between logic, law and history?
3. Demonstrate the use of deductive reasoning in application of a statute and/or the use of inductive reasoning in a judgment of case law.

Additional resources

For examples of derivation of a general principle through inductive generalisation, see Brett MR's judgment in *Heaven v Pender* (1883) 11

QBD 503; Lord Atkin's judgment in *Donoghue v Stevenson*; Blackburn J's judgment in *Fletcher v Rylands* (1866) LR 1 Exch 265; Lord Denning's judgment in *Lloyd's Bank v Bundy* (1975) QB 326.

In relation to analogical reasoning, see Goddard LJ's judgment (and Clauson LJ's dissent) in *Haseldine v Daw* (1941) 22 KB 343; for an example of distinguishing a precedent on grounds of relevant disanalogy, see Lord Denning's judgment in *Thornton v Shoe Lane Parking Ltd* (1971) 2 QB 163.

Chapter 3

FORMAL FALLACIES AND FALLACIES OF RELEVANCE

Fallacies

Logic is just as much concerned with bad reasoning as with good. In particular, logicians have identified a range of common errors in reasoning, which go beyond mere falsity of premises. Specifically, they have identified a range of argument forms that have the appearance of being deductively sound or inductively strong (or cogent), when, in fact, they are not. Such arguments, which mislead people into thinking that they are good when they are not, are called *fallacies*.

Fallacies can be the results of simple errors in reasoning. And here it seems that humans are particularly prone to particular sorts of errors, perhaps by virtue of deep structures of their brains, or deep structural features of their social situation.

But fallacies can also be constructed for the express purpose of producing particular results. On the one hand, individuals can construct fallacious arguments (while unaware that they are fallacious) for the purpose of convincing themselves of the truth of appealing (but unsupported) ideas and the falsity of unappealing (but strongly supported) ones.

On the other hand, individuals can quite consciously produce and disseminate fallacies with a view to fooling and manipulating other people seeking to establish in them beliefs that support the interests of the ideas producers or their patrons, but that have no solid basis in fact or moral principle.

Advertisers, political parties and corporate-sponsored 'think tanks' such as 'Consumer Alert', 'The National Wetlands Coalition' and the

'Tobacco Industry Research Council' in the United States and similar organisations in Australia have devoted vast time and resources to such systematic manipulation through construction and propagation of fallacious arguments. As noted earlier, lawyers have all too often been guilty in this regard also.

The richest and most powerful lawyers are rich and powerful because they represent the richest and most powerful organisations and individuals, which, in capitalist society, are mainly major corporations, their managers and leading shareholders (but also include politicians, entertainers and sports stars). Many such lawyers work in the inhouse legal departments of the corporations themselves as employees. Others work in huge law firms specialising in corporate clientele. Vast sums of money are involved here with the largest grossing law firms in the United States producing billions of dollars of revenue each year, and leading corporate lawyers, particularly corporate employees, earning multimillion dollar salaries. Even beginning lawyers hired by big corporate law firms earn far in excess of average wages and gain numerous special benefits.[1]

Such lawyers are involved in vigorous competition among themselves for partnership positions 'where the real influence and money awaits'.[2] And as Nader and Smith argue, 'in this intensely competitive atmosphere, idealism and rigorous adherence to professional ethics can easily take a backseat to simply getting the job done'.[3] They quote Michael Josephson, a lawyer running a California-based ethics institute, to the effect that

> Lawyers are competitors. They tend to look at what they do as a sport. The way law is practised today is to get away with what you can. Why? Because money is the only way to judge … The big firms are doing the bad things because the big firms have the big cases.[4]

Attorney Steve France, editor of *Bank Lawyer Liability*, notes that business executives

> want lawyers with guts of steel and no morals. Under this theory lawyers are machines – you just point them in the direction you want and away they go. This stick-it-to-the-other-guy approach to law, especially if the other guy is weaker and perceived to have fewer resources, creates an attitude that often condones misbehaviour and thereby extracts a terrible price, often in the lies of ordinary people made unjustly miserable by the blitzkrieg methods of power lawyers.[5]

Nader and Smith conclude:

> This is cause for alarm. Unless tempered by adherence to the higher calling of professional honour and restraint, unquestioning client loyalty can cause profoundly adverse consequences. The lawyer devolves into the proverbial hired gun, where, as in the old Westerns, an ethic of might makes right prevails. Reputations are made, not by doing right, not by avoiding wrongdoing, not by wise counsel, but by winning. The system descends into legal Darwinism, as natural selection increasingly favours those clients with the most money to spend on lawyers willing to do whatever it takes to please the patron.[6]

For present purposes, the crucial point is that such 'guns for hire' are unlikely to feel any compunction in setting out to mislead and manipulate judge and jury through active use of fallacious argument. As suggested earlier, judges intent upon realising particular political agendas, without acknowledging that this is what they are doing, can also utilise fallacious reasoning to mask their actions. They can offer logical justifications that are really rationalisations for their manipulation of the law in one direction or another.

Fallacies of all kinds can be extremely dangerous, whether individuals fool themselves or allow others to fool them. And it is crucial that individuals sharpen up their abilities to detect and refute fallacies, if they are not to be regularly misled, manipulated and exploited. If law students are genuinely interested in justice, then it is crucial that they are ready and able to identify and unmask the fallacies of less scrupulous practitioners and lawmakers.

Formal fallacies

Formal fallacies can be identified through consideration of the structure of the argument. In particular, people seem to be prone to confuse necessary and sufficient conditions in conditional statements leading to the fallacy of *affirming the consequent*. Here, the truth of the consequent is taken as sufficient for the truth of the antecedent, when it is really only necessary.

P1 If it rains, the car will get wet at that time.

P2 The car is wet now.

C Therefore, it is raining now.

The truth of the first two statements might provide some inductive evidence for the truth of the third. But if this is presented as a valid deductive argument, where the conclusion follows necessarily from

the premises, as it appears to be, then it is clearly a fallacy. There are lots of other ways the car could have got wet. Rain is sufficient to make the car wet, but not necessary.

A closely related fallacy is that of *denying the antecedent*:

P1 If someone drives in the park then they break a rule.

P2 Andrea didn't drive in the park.

C Therefore, Andrea didn't break a rule.

Again, the premises might provide some (very weak) inductive support for the conclusion. But the conclusion certainly does not follow with deductive necessity. Andrea might not have broken the rule about driving in the park. But she has just murdered her employer and has his body in the boot of her car along with boxes of heroin and unlicensed guns. She has broken a number of other rules. It is sufficient to drive in the park to break a rule. But it is not necessary. There are lots of other ways to break rules.[7]

There is a well-known psychological test in which subjects are presented with four cards, one displaying a vowel, one a consonant, one an even number and one an odd number (for example, A R 2 7). Subjects are told that each card has a letter on one side and a number on the other. They are asked which cards they need to turn over to test the hypothesis that if a card has a vowel on one side then it has an even number on the other. Most subjects turn over the vowel card and the even number card. Very few turn over the odd number card. These results are commonly taken to show a 'confirmation bias': that is, a tendency to look for confirming rather than refuting evidence. If the vowel card does have an even number on the other side, the hypothesis is confirmed, similarly if the even number has a vowel on the other side.

Certainly, these results suggest a general failure to grasp the relations of necessary and sufficient conditions underlying the deductive arguments of *modus ponens* and *modus tollens*. To see this we identify vowel = p, consonant = not p, even number = q, and odd number = not q. As we know, the hypothetical statement 'if p then q' says that p is sufficient for q, q is necessary for p. As we also know, 'if p then q' is quite a different statement from 'if q then p'. Finding a vowel behind the even number does provide some inductive support for 'if p then q', but finding a non-vowel does not refute the hypothesis.

If it were costly to turn over the cards we should probably ignore the even number card. On the other hand, because q is necessary for

p in 'if p then q', then 'if p then q' is logically equivalent to 'if not q then not p'. So we should turn over the odd number (not q). If we find a vowel behind it we have refuted the hypothesis. Similarly, we should turn over the vowel card, not primarily because of the opportunity of providing some inductive support for the hypothesis, but rather because of the possibility of definitive refutation – if there is an odd number on the other side.

Fallacies of relevance

Informal fallacies are those that can be detected only through analysis of the content of the argument. There are many sorts of informal fallacies, which can be classified in different ways. Some particularly important types have been called fallacies of relevance, weak induction, presumption, ambiguity, and grammatical analogy. But such categories cannot be rigidly demarcated and inevitably overlap at various points.

Fallacies of relevance substitute psychologically or emotionally appealing connections between premises and conclusions for logically relevant connections, so that the conclusion can seem to follow when really it does not.

Appeals to force suggest that listeners should accept certain conclusions to avoid some harm that will come to them if they do not. *Appeals to pity* aim to get listeners to accept particular conclusions out of pity for some individual or group. *Appeals to the people* (or *ad populum* arguments) play upon an individual's need to be loved, esteemed, admired and accepted by others.[8]

We might imagine all of these fallacies intruding into protracted jury deliberations, particularly where one juror holds out in face of unanimity among the rest.

P1 You want to get out of here before the weekend and see your family/you want to be appreciated and approved by the rest of us jurors and all right thinking people/you want to be kind to the family of the victim.

P2 You can do these things by agreeing with the rest of us that the accused is guilty.

C Therefore you should agree that they are guilty.

The *straw man fallacy* is committed when the arguer distorts an opponent's argument for the purpose of more easily attacking it, refutes the distorted argument, and then concludes that the opponent's original

argument has been refuted.[9] This sort of fallacy is particularly common in political debate, and particularly in relation to proposals for reform of current institutions and practices. Defenders of the status quo are typically reluctant to seriously engage with any such proposals and tend, instead, to fall back upon the creation of straw men. For example, a critic of 'anti-terrorism' legislation might be accused of supporting terrorism.

The adversarial format of the common law also provides strong temptations to misrepresent an opponent's argument in such a way as to make it easier to attack it. We are all familiar with such a process in a legal context from numerous fictional portrayals of lawyers 'summing up' their cases, even if we have no experience of real court cases. Judges and juries have to be on their logical toes to recognise that this is, indeed, happening.

Ad hominem arguments direct attention to the (alleged) bad character of a person involved in presenting their own logical argument, and away from the real structure and content of that person's argument. The arguer themselves is criticised or denigrated, with the aim of getting some third party to see the limitations of the arguer as limitations of their argument.

P1 Smith says heroin should be legalised.

P2 Smith is a communist/drug addict/prostitute.

C Therefore her argument is no good and we should reject it.

This fallacy also is common in political debate, particularly in relation to radical critique.

P1 Karl Marx lived off the profits of Frederick Engels' cotton mills.

P2 He failed to acknowledge paternity of an illegitimate son.

C Therefore we cannot take seriously any of his critical ideas about capitalist society.

In a criminal legal context we can see how jurors might be deflected from serious consideration of the facts of the current case by evidence of the previous 'bad character' of the accused, particularly any record of previous criminal acts. The courts have taken steps to try to avoid this.

> In the case of *R v Perrier* (No 1) [1991] 1 VR 697, the Supreme Court of Victoria made it clear that normally the prosecution cannot lead evidence of an accused's bad character. It is only if the accused has

> attempted to establish her or his own good character [and hence put that character at issue] that the evidence of the accused's bad character will be admissible. The court also stressed that it is the conduct of the accused in raising the issue of his or her own character that makes the evidence of bad character admissible.[10]

This 'character shield' is formalised in *Evidence Act 1995* (NSW), ss 110 and 104. Section 104 states:

> Leave must not be given for cross-examination by the prosecutor about any matter that is relevant only because it is relevant to the defendant's credibility unless (a) evidence has been adduced by the defendant that tends to prove that the defendant is, either generally or in a particular respect, a person of good character, or (b) evidence adduced by the defendant has been admitted that tends to prove that a witness called by the prosecutor has a tendency to be untruthful, and that is relevant solely or mainly to the witness's credibility.

It is good that the accused cannot, as a matter of course, be forced to reveal facts about themselves that might prejudice the jury against proper consideration of the facts of the present case. But it is unfortunate if defendants, in attacking the reliability of the prosecution witnesses, thereby open themselves to what may prove to be far more damaging revelations about themselves, which do indeed encourage fallacious reasoning on the part of judge or jury.

The fallacies of *missing the point* (*ignoratio elenchi*) and *red herring* (distracting the hunting dogs with its smell) are often difficult to clearly distinguish in practice. In the former case, the premises support one conclusion, but the arguer, instead, draws a different one, only vaguely related to the correct one. In the latter,

> the arguer diverts the attention of the reader or listener by changing the subject to some totally different issue. He or she then finishes by either drawing a conclusion about this different issue or by merely presuming that some conclusion has been established.[11]

Gigerenzer identifies a particularly insidious example of a red herring in the OJ Simpson criminal case. As he says:

> Alan Dershowitz, a renowned Harvard law professor, advised the Simpson defence team. In his best-selling book *Reasonable Doubts: The Criminal Justice System and the OJ Simpson case*, Dershowitz explained the team's success in quashing the prosecution's argument that spousal abuse and battering should not be admissible in a murder trial: 'The reality is that a majority of women who are killed are killed by men with whom they have a relationship, regardless of whether

> their men previously battered them. Battery, as such, is not a good independent predictor of murder.'
>
> The defence told the court that some studies estimated that as many as 4 million women are battered annually by husbands and boyfriends [in the United States]. Yet in 1992, a total of 913 women were killed by their husbands, and 519 … by their boyfriends … From these figures Dershowitz calculated that there is less than 1 homicide per 2500 incidents of abuse … [W]e were convinced from the very beginning that the prosecutors' emphasis on what they called 'domestic violence' was a show of weakness. We knew that we could prove … that an infinitesimal percentage … of men who slap or beat their domestic partners go on to murder them.' Dershowitz concluded 'there is never any justification for domestic violence. But neither is there any scientifically accepted evidence that domestic abuse … is a prelude to murder'.[12]

As Gigerenzer points out, the figure of one homicide per 2500 incidents of abuse is largely irrelevant in determining the likely guilt of OJ Simpson, or any abuser accused of murdering their abused partner. As he says, the relevant percentage is not how many men who slap or beat their domestic partners go on to murder them, [rather] the relevant probability is that of a man murdering his domestic partner given that he battered her and that she was murdered.'[13]

Using figures from the *Uniform Crime Reports for the United States and its Possessions* (1993) Gigerenzer shows that around 40 out of every 45 murdered and battered women are killed by their batterers every year in the United States. 'That is, in only 1 out of 9 cases is the murderer someone other than the batterer.'[14] Far from being irrelevant to the issue of his guilt, as Dershowitz argues, the fact that OJ Simpson battered his wife was profoundly relevant to the question of his guilt in relation to her murder. Gigerenzer concludes:

> This probability [the 8 out of 9] must not be confused with the probability that OJ Simpson is guilty; a jury must take into account much more evidence than battery to convict him beyond a reasonable doubt. But this probability shows that battering is a fairly good predictor of guilt of murder, contrary to Dershowitz's assertions. Evidence of battery is probative not prejudicial.[15]

The prosecutor's fallacy

The best-known example of systematic fallacious reasoning (in the red herring category) in a legal context is that of the *prosecutor's fallacy*. This has become particularly significant with the advent of DNA fingerprinting, and has been responsible for major miscarriages of justice.

> The term seems first to have been used by attorney and social psychologist William Thompson and his student Edward Schumann in relation to a deputy district attorney's argument that if a defendant and perpetrator match on a blood type found in 10% of the population, there is a 10% probability that the defendant would have the blood type if innocent and therefore a 90% probability that they are guilty.[16]

The 10% figure is called a *random match probability*: in other words, the chance that a randomly selected individual will exhibit the property in question. There is no way a probability of guilt can be derived from any such figure. On its own it is indeed a red herring in relation to any attempt to establish likelihood of guilt.

We need, at a minimum, to have also information about the pool of possible perpetrators – about those who could, possibly, have committed the crime. Assuming that anyone in a population of 100 000 people could have done it, then the matching blood types narrow the figure down to the 10% of that number with the blood type in question – 10 000 people. The match indicates a 1-in-10 000 chance of guilt, rather than 90%.

The prosecutor's fallacy confuses the probability of innocence given a match (the figure of interest to judge and jury) with that of a match given innocence (the random match probability). That these are far from being the same thing is shown by examples such as the following: The probability of being a head of state or CEO of a major corporation given that you are a man is virtually zero. The probability that you are a man given that you are a head of state or CEO is virtually 100%.

If you are innocent, the probability that your blood type matches that at the scene of the crime is the random match probability (in this case, 10%). But, as noted earlier, (a rational assessment of) the probability of innocence given a match depends also upon the population of possible perpetrators and upon any other relevant evidence that might be available.

In fact, in all cases where bodily tissues at the scene of the crime are an issue, there are a number of things to consider:

> First, a reported match [in tissue types] may not be a true match because of laboratory errors … that produce false positives.[17]

So we need to consider the 'error rate' of the laboratory concerned.

> Second, a defendant who provides a true match may not be the source of the trace if the match is co-incidental; even rare tissue types can occur in more than one person, particularly in biological relatives.[18]

> Third, a defendant who is truly the source of the trace may not have been present at the crime scene, if … someone else deliberately, or unintentionally … transferred the defendant's biological material to the scene … [And] finally, a defendant who had been present at the crime scene may not be guilty – they may have left the trace before or after the crime was committed, or been an innocent bystander.[19]

Bayes' theorem and prior probabilities

Another variation of the prosecutor's fallacy, or of the broader category of ignoring prior or unconditional probabilities, is nicely illustrated by reference to a test problem used by psychologists David Kahneman and Amos Tversky:[20]

> A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. The following facts are known:
>
> 85% of the cabs in the city are green and 15% are blue,
>
> A witness identified the cab as blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colours 80% of the time and failed 20% of the time.
>
> What is the probability that the cab involved in the accident was actually blue?[21]

The typical answer, from subjects in all walks of life, including lawyers, was around 80%. People think 'it does not matter what percentage of cabs in the city are blue. What matters is the eyewitness testimony. The eyewitness is right 80% of the time and they say it is blue. So it probably is blue'. The 'prior' or unconditional probability of its being blue is ignored. But it shouldn't be.

To find the correct answer we can employ a mathematical formula called Bayes' theorem to the effect that Pr(A/B) = Pr(B/A) x Pr(A)/Pr(B). Pr(A/B) is the probability of A given B; the cab is blue given that the witness says it is. This is what the court needs to know. Pr(B/A) is the probability of B given A: the witness says it is blue, given that it is blue. This probability is 80%; if the cab is blue the witness will correctly identify it as blue 80% of the time. This is analogous to the random match probability shown in the earlier prosecutor's fallacy example. And, as in that case, the subjects of the test have confused this figure – on its own irrelevant to issues of guilt and innocence – with the figure of real significance to judge and jury, the probability that the cab really was blue, given the witness says it was.

Bayes' theorem enables us to find the probability it is blue, given the witness says it is (the figure the court should be interested in) by considering the probability of the witness saying it is blue, given that it is blue. But to do so, we need to also consider Pr(A) and Pr(B). Pr(A) is the unconditional or prior probability that the cab is blue, which is 15% (0.15). Pr(B) is the unconditional probability that the witness says it is blue. This is the tricky bit.

> Because the witness is not infallible, they will on some occasions correctly identify a blue cab and on others incorrectly identify a green cab as blue. [And because there are a lot of green cabs they will call quite a few green cabs blue]. Consequently the required unconditional probability is given by the sum of the probabilities of these two events:
>
> Cab is blue and witness correctly identifies it as blue [0.15 x 0.80].
>
> Cab is green and witness incorrectly identifies it as blue [0.85 x 0.20].
>
> Therefore the required unconditional probability of the witness saying the cab is blue is [0.15 x 0.80] + [0.85 x 0.20] = 0.29.[22]

We now have all that is necessary to apply Bayes' theorem, which yields: Pr(A/B) = 0.80 x 0.15/0.29 = 0.41 (approximately).

Most people, including lawyers, thought the cab was most probably blue. But in fact, there is less than a 50% chance that it was blue; it was probably green. This is certainly an artificial example. But it is not a million miles from the sort of consideration with which real judges and juries are confronted every day.

A somewhat bizarre finding in the Court of Appeal in the United Kingdom in 1997 stated that 'introducing Bayes' theorem … into a criminal trial … plunges the jury into inappropriate and unnecessary realms of confusion' and Bayes should therefore be excluded from future criminal proceedings. Nonetheless, informed members of the legal profession now recognise that DNA evidence, in particular, has to be understood in terms of Bayes' theorem in order to yield any genuinely legally significant results.[23]

Although the Court of Appeal was mistaken in arguing that Bayes should be kept out of the courtroom, it is true that the mathematics remains intimidating to many people. But Gigerenzer has shown how the same information can be much more accessibly presented in diagram form in terms of 'natural frequencies'.

If we take the cab case and consider 1000 cabs in the city, 150 of them are blue (15%) and 850 are green (85%). The witness sees 120 of the blue cabs as blue (80%). Thirty are seen as green (20%). She

sees 680 of the green cabs as green (80%) and 170 as blue (20%). So the total number seen as blue = (120 + 170) = 290. So Pr(actually blue/identified as blue) = 120/(120 + 170) = 120/290 = 12/29 = 0.414%.[24]

Discussion topics

1. What is a fallacy? Include reference to formal fallacies and fallacies of relevance.
2. What is the prosecutor's fallacy? Include reference to Bayes' theorem and forensic DNA testing.
3. Find and explain some examples of fallacious reasoning from textbooks, newspapers and judgments.

Additional resources

B. Atchison, 'Criminal Law: DNA Statistics May Be Misleading', *Law Society Journal*, February 2003, p 68.

Chapter 4

WEAK INDUCTION AND PRESUMPTION

In the previous chapter we considered formal fallacies and fallacies of relevance. There is a wider range of informal fallacies. Among the most important are those relating to weak induction. These include weak analogies and hasty inductive generalisations.

Weak analogy

As Hurley says:

> The fallacy of *weak analogy* is committed when the analogy is not strong enough to support the conclusion that is drawn. The basic structure of an argument from analogy is as follows:
>
> P1 Entity A has attributes a, b, c, and z.
>
> P2 Entity B has attributes a, b, c.
>
> C Therefore, entity B probably has attribute z also.[1]

If there is some 'causal or systematic relation' between z and a, b, and c, and this is not 'offset' by strong disanalogy, the argument is strong. But if there is not then it is weak.

It is less than immediately obvious what sorts of connections are likely to be relevant in a legal context. Different sorts of connections seem to be relevant in different legal situations. But it is certainly possible to detect weak analogies in legal reasoning.

It is easy to see how stare decisis might lend itself to weak analogising as a rationalisation for politically motivated decisions that might be difficult to defend as such (claiming relevant similarities with

a prior case where the 'desired' result was achieved, where no such relevant similarities actually exist, or where they are more than offset by relevant dissimilarities).

One example is found in *Hill v Chief Constable of West Yorkshire* (1989) AC 53. The case arose out of a series of horrific murders, in Leeds, Bradford, Halifax, Huddersfield, Farsley and Manchester in northern England. Over a five-year period from October 1975 to November 1980, Peter Sutcliffe, the Yorkshire Ripper, killed 13 women in northern England. He bashed their heads with a hammer and mutilated their bodies.

The mother of Sutcliffe's last victim sued the Chief Constable of West Yorkshire for negligence on behalf of her daughter's estate. It was claimed that during the police investigation of the series of crimes a number of mistakes were made such as the failure to compare and evaluate properly information in possession of the police pointing to the perpetrator. It had to be assumed that he would have been arrested before the murder of the plaintiff's daughter if the police had exercised reasonable care and skill.

The case followed two others in which the plaintiff attempted to sue the Chief Constable of West Yorkshire as vicariously liable for the torts of police officers under his direction, under the *Police Act 1964* (UK). In both earlier trials the statement of claim was struck out as 'disclosing no reasonable cause of action'. But the plaintiff pursued the matter in the House of Lords.

In the leading judgment (with which the other judges agreed), Lord Keith focused mainly upon the prior case of *Dorset Yacht Co Ltd v Home Office* (1970) AC 1004, as establishing an appropriate precedent, insofar as it involved 'reasonable foreseeability of likely harm arising from failure to control another man to prevent his doing harm to a third'.[2]

In *Dorset Yacht*, owners of yachts damaged by Borstal escapees under the care of prison officers successfully sued the Home Office for compensation. The Home Office argued that Borstal officers owed a duty only to the Crown, not to the public, with no previous precedents for such Home Office employees or employees of other statutory authorities owing any such duties to the public. But the Appeal Court rejected this argument, referring instead to the neighbour principle to find that such officers did indeed owe a duty of care to certain members of the public under certain circumstances. The neighbour principle should be generally applied, 'unless there were some good reasons to the contrary'.

The court spent some time in considering issues of 'remoteness' in limiting (or clearly defining) the application of the principle in 'this sort of case.' Lord Diplock compared the specific circumstances of *Dorset Yacht*, involving criminal damage by known offenders inflicted upon property very close to their area of confinement in the course of their escape with a hypothetical alternative case. In *Hill*, Lord Keith summarised what he took to be Lord Diplock's position, to the effect that

> no liability would rest upon a prison authority which carelessly allowed the escape of a habitual criminal, for damage which he subsequently caused, not in the course of attempting to make good his getaway to persons at special risk [as with the boat-owners in this case] but in further pursuance of his general criminal career to a person or property of members of the general public.[3]

Lord Keith interpreted this judgment to have established that a specific sort of 'proximity', determined by the 'category of case', was now a necessary condition of liability in negligence. In *Dorset Yacht*, the warders knew of the identity and criminal proclivities of the escapees prior to their escape, they should also have seen that the yachts were at special risk from such escapees, by virtue of their location, and they could have taken speedy action to prevent the damage without radical disruption of their 'normal' responsibilities, by virtue of their close physical proximity to the yachts in question.

The suggestion is that others would be liable in negligence to third party victims only if those others were (similarly) aware of the identity and criminal habits of relevant second parties, and only if the relevant victims could reasonably have been identified (by those others) as at special risk prior to the crimes in question.

This seems reasonable enough – at least at first sight – in relation to prison warders. Once an escaped criminal has passed out of physical proximity to the institutional base of the warders – with no particular 'person' (or property) any longer identifiable as at special risk – such escapees should be taken as having passed beyond such warders' legal responsibility of recapture or protection of potential victims. Warders lack resources and training to go chasing around the country after escaped prisoners or protecting the general public from the criminal actions of such ex-prisoners. Certainly they lack resources to determine the identities of unknown criminals. On the contrary, this would seriously interfere with their primary responsibilities of care in relation to fixed prison sites if they abandoned such sites in favour of hunting down escaped career criminals.

But then Lord Keith proceeded to apply these principles in the *Hill* case. As he said, the police lacked the appropriate relationship of proximity to both killer and victim. They did not know the killer's identity or habits prior to the murder, and they could not reasonably have identified the victim as someone at special risk.

> Sutcliffe was never in the custody of the police force. Miss Hill was one of a vast number of the female general public who might be at risk from his activities but was at no special distinctive risk in relation to them.[4]

Several problems arise here. The whole point of Mrs Hill's argument was, of course, that the facts show that the police should have known the identity of the killer and/or victim prior to the murder, had they done their job properly. But the Law Lords did not consider these facts and therefore entirely 'begged the question' really at issue.

Lord Keith defended his claim of no special risk to Miss Hill simply by saying 'all householders are potential victims of an habitual burglar'.[5] From this we are presumably supposed to infer that all women are potential victims of an habitual woman-killer.

Lord Keith's premise is simply false. Particular burglars are only able, or likely, to target particular households. Greater knowledge of relevant criminological data and the empirical details of particular cases could be expected to significantly narrow the focus of relevant investigation. And the same applies in relation to the much less common crime of serial murder.[6]

Most important, it is not reasonable to apply principles derived from consideration of the responsibilities and practical capacities of prison warders to the police force. There are significant disanalogies between police and prison officers. The very things which it is unreasonable to expect prison warders to do are precisely the things that police are employed – and amply funded – to do, due to a clear division of labour between agencies of criminal law enforcement. Prison warders (supposedly) protect the public through keeping known criminals in jail (perhaps even trying to rehabilitate them); police protect the public by identifying and apprehending previously unknown criminals. And proper investigations (of previous crimes) might, in some cases, at least, be expected to allow for identification and protection of likely future victims. Certainly, we should expect the police to make reasonable efforts in these areas – and hold them responsible where their negligence leads to disastrous results.

It seems, indeed, as if the Law Lords, led by Lord Keith, were determined to radically limit police liability in negligence, and were prepared to use the weakest of arguments – particularly the 'weak analogy' comparing the police with prison officers – to achieve this.

Hasty generalisations and false causes

> Inductive generalisation involves the drawing of a conclusion about all members of a group from evidence that pertains to a selected sample. The fallacy [of *hasty generalisation*] occurs when there is a reasonable likelihood that the sample is not representative of the group. Such a likelihood may arise if the sample is either too small or not randomly selected.[7]

Hasty generalisations in this sense are very common and form the basis for much stereotyping and prejudice. On the basis of limited experience of particular ethnic or cultural groups (or on the basis of a very biased selection of people or events), some people are ready to make sweeping general statements about the nature of all members of such groups – unsupported by the facts available to them. People see only the crimes committed by particular members of such groups and leap to the conclusion that all members of such groups are criminals, often due to systematic scapegoating by mass media outlets.

Such prejudices can be particularly dangerous when established within agencies of law enforcement, where they can motivate and legitimate persecution and victimisation of particular groups by armed, intimidating and dangerous police officers. And, unfortunately, there is substantial empirical evidence of just such prejudices firmly established amongst such agencies around the world.[8]

We earlier noted Lord Atkin's argument that Brett MR's inductive generalisation in relation to duty of care in negligence was too hasty in this sense. Others, including the dissenters at the time, have seen Lord Atkin's own inductive leap to the neighbour principle as similarly unjustified.

A number of different fallacies involve *causation*. Hurley identifies the best-known:

> The fallacy of *false cause* occurs whenever the link between premises and conclusion depends upon some imagined causal connection that probably does not exist.[9]

> The fallacy of *slippery slope* is a variety of the false cause fallacy. It occurs when the conclusion of an argument rests upon an alleged chain reaction and there is not sufficient reason to think that the chain reaction will actually take place.[10]

We earlier referred to the institutionalisation of fallacious argument in the common law. It was the slippery slope fallacy, in particular, that we had in mind. This sort of argument crops up again and again, both in relation to law reform and, particularly, in so-called policy arguments by judges in the higher courts.

Proposals for, or practices of, legalisation of active, voluntary euthanasia often provoke comparison with the Nazi Holocaust, in which euthanasia of the mentally and physically damaged ultimately extended into mass genocide. It is suggested that active euthanasia contributes to a disrespect for the value of human life that leads, ultimately, to mass murder. Here, faulty analogy jostles with false cause and slippery slope. More temperate critics argue that even attempts to clearly confine euthanasia and/or assisted suicide to fully informed and voluntary decisions of terminally ill people will inevitably fail, leading to less than fully voluntary or informed termination of those not necessarily terminally ill.

A similar situation exists in relation to the duty of rescue. Unlike continental civil law, the common law recognises no general duty to rescue others in distress, even when this can be accomplished with ease and little threat to the rescuer: that is, where great benefit can be achieved, great harm can be avoided, at little cost. A frequent argument is that once we start down the path of liability for omissions in this area, there will be no turning back, with those who fail to give alms to a beggar held responsible for the beggar's subsequent illness or death, perhaps years later.

One obvious counter-argument in relation to rescue is that the line should be drawn at clear cases of opportunity to rescue others from serious and immediate distress, with no unreasonable risk, cost or inconvenience, with the definition of reasonableness left to prosecutors, judges, juries and legislators, as in other areas of law. Another would refer to the operation of the principle in civil law jurisdictions.

Slippery slopes have also been prominent in the increasingly prevalent policy arguments of presiding judges, pushing the law in particular directions. Cook et al note:

> In judicial decision-making at the appellate level, principle and policy play an increasingly important – some would say a dominant – part, in addition to legal authority or the doctrine of stare decisis.[11]

A large number of so-called policy decisions revolve around what has been called the 'floodgates' argument. Such arguments assert a narrowly

circumscribed formulation of a principle seen to underlie a group of previously decided cases and then argue that any broadening of such a principle will inevitably create the grounds for a vast increase in new actions, lacking in real moral foundation and/or threatening to completely overwhelm the legal system itself, and/or leading to other highly socially undesirable consequences (of collapse of civilisation as we know it).

It is true, as MacAdam and Pike point out, that in many milestone torts cases since *Donoghue v Stevenson*, 'the floodgates argument has received pretty short shrift'.[12] But arguments of this form have still been influential in shaping the recent development of law in a number of crucial areas. And as MacAdam and Pike admit, 'the main function of policy arguments in negligence cases is as a factor limiting the breadth of application of the neighbour principle'.[13]

Looking back to the *Hill* case, already considered in relation to faulty analogy, we see such a policy argument aiming to justify the radical restriction of police liability in negligence by reference to a (problematic) slippery slope. Thus, at the end of his judgment, Lord Keith says:

> Potential existence of liability [in negligence] may in many instances be in the general public interest, as tending towards the observance of a higher standard of care in the carrying out of various different types of activities. I do not however consider that this can be said of police activities. The general sense of public duty which motivates police forces is unlikely to be appreciably reinforced by the imposition of such liability so far as concerns their function in the investigation and suppression of crime. From time to time they make mistakes in the exercise of that function, but it is not to be doubted that they apply their best endeavours to the performance of it. In some instances the imposition of liability may lead to the exercise of that function being carried out in detrimentally defensive frame of mind.[14]

Furthermore, 'preparation of the defence' of such actions would result in a

> significant diversion of police manpower and attention from their most important function; that of the suppression of crime … Closed investigations would require to be re-opened and re-traversed, not with the object of bringing any criminal to justice but to ascertain whether or not they have been competently conducted.[15]

The problems here are by no means confined to the slippery slope. What is perhaps most striking is the way in which the judge seems to take for granted that he can present sweeping empirical generalisations with absolutely no empirical evidence presented to support

them. And this is especially so given the substantial body of data that clearly refutes his assertions, including numerous police royal commissions and similar investigations around the world that consistently show the police to be no more free from corruption and incompetence than other professions.[16]

In answer to the final slippery slope suggestion, that is, that liability in negligence will radically undermine police effectiveness in 'investigation and suppression of crime' through increasing 'defensive' tactics and greater and greater time taken up in 'preparing defences', we can refer to the case of other professions where this has not happened and, more relevantly, to other jurisdictions that have held the police responsible in negligence in similar circumstances (for example, Canada and Germany)[17] without any such consequences.

The 'other side' of the false cause fallacy (identifying a non-cause as a cause) is the fallacy of denying that a real cause actually is a cause or *cause as non-cause*. It is easy to see the importance of this in negligence cases where a great deal can hang upon the issue of causation and the relevant (scientific) evidence can be quite complex.

Thus, the plaintiff might advance scientific evidence to demonstrate the likelihood that their illness or injury (their lung cancer, for example) has been produced by the defendant's product or by-product (cigarettes or asbestos or benzene or diesel fumes), and that the defendant had or should have had knowledge of the likelihood of such an eventuality (relevant evidence was readily available in the public domain prior to the exposure in question) and has therefore been negligent in producing and selling the product in question (with inevitable exposure of workers or consumers).

Counsel for the defendant might reply that the evidence cited by the plaintiff does not actually establish causation at all, but merely demonstrates a 'statistical' relationship, or a correlation between the product and the injury or illness in question. The mere fact that significantly more smokers than non-smokers develop lung cancer proves nothing about causation: perhaps (a very early stage of development of) lung cancer causes smoking, perhaps some third (possibly genetic) factor causes both. This line is all too familiar from decades of arguments about tobacco products.

Correlation is indeed not the same as causation. But correlation can provide (indirect but strong) evidence of causation. If we can precisely map the incidence of an extremely rare cancer to sales of a par-

ticular drug 20 years previously, for example, then we have a correlation strongly suggestive of causation.[18]

Evidence of causation can be 'statistical' in the sense that we know only what percentage of cases of particular illnesses are caused by exposure to particular toxins or carcinogens, without knowing which particular cases are so caused. But with use of appropriate control and experimental groups and statistically significant results it can be strong evidence of causation nonetheless.

The law has been slow to come to terms with modern scientific understanding of the nature of causation. This issue is considered in detail in the next two chapters.

Argument from authority

The 'other side' of *ad hominem argument* is *argument from authority* (*ad Verecundiam*). In this case, the argument is uncritically accepted on the basis of the authority of the arguer, without serious consideration of its real, logical strength.

The situation here is complicated by the fact that appropriately qualified authorities generally can be – and sometimes have to be – trusted to have good reasons for considered pronouncements within their area of special expertise. As Fogelin and Sinnot-Armstrong note:

> An authority is a person or institution with a privileged position concerning certain information … [A]n appeal to experts and authorities is essential if we are to make up our minds on subjects outside our own range of competence.[19]

Thus, the *ad Verecundiam* is sometimes equated only with 'Appeal to Unqualified Authority'. Certainly, we can see the danger in trusting an expert in area A (for example, medicine) to make reliable judgments in a quite different area B (for example, nuclear physics).

Here again, rules of evidence are supposed to guard against this sort of thing. As Dawson J noted (in *Murphy v R* (1989) 167 CLR at 131):

> [E]ven though most juries are not prone to pay undue deference to expert opinion, there is at least a danger that the manner of its presentation may, if it is wrongly admitted, give to it an authority which is not warranted.

Thus, in Australia:

> There must be a correlation between the field of knowledge about which a witness is to testify and that witness' qualifications and

> experience. A trial judge must ascertain this correlation by identifying the issue about which the witness is to testify and comparing this issue with the witness' qualifications and experience.[20]

But even if we are concerned with judgments within the area of the expert's special skill and knowledge, it is still dangerous for jurors or anyone else to merely 'accept' the truth of such pronouncements without critical consideration of the underlying evidence and argument. If the alleged expert really is an expert in the area concerned they should be able to clearly explain the reasoning involved.
Again, rules of evidence require that

> In order to admit expert opinion, it is necessary for the expert to demonstrate the assumptions, or the facts, upon which his or her opinion is based …[21] The failure of an expert witness to state the assumptions upon which her or his opinion is based is likely to render the evidence of the expert inadmissible.[22]

Nonetheless, the special position of the expert – or alleged expert[23] – still gives them special opportunity to mislead judge and jury. And it is important not to neglect the forces militating towards such a development in the modern world. In particular, we must remember that many scientists are now employees of big private corporations, or are themselves corporate entrepreneurs, in both cases committed to maximisation of short-term profits, rather than to any idea of the public good.

This does not, of course, necessitate lying and cheating on the part of these employees (in the service of corporate profits) in their research publications or their public statements. But it can mean this in some cases. And we must not underestimate the power of the employment contract or research funding, in a time of widespread unemployment and increasing reduction in public research funding, to even unconsciously influence the behaviour of a scientific expert witness in a court room situation, or out of it.

Fallacies of presumption

In some arguments the premises presume what they purport to prove. As Hurley says:

> *Begging the question* occurs when the arguer uses some form of phraseology that tends to conceal the questionably true character of a key premise. If the reader or listener is deceived into thinking that the key premise is true, he or she will accept the argument as sound, when in fact it may not be.[24]

We have already seen an example of begging the question in the *Hill* case. As noted, Lord Keith simply stated that the police did not know the identity of the killer at the time of the murder of Mrs Hill's daughter. This was then used, by reference to *Dorset Yacht* as precedent, to exempt the police from negligence in relation to that murder. But the point of Mrs Hill's argument was that the facts showed that the police should have known the identity of the killer and/or victim prior to the murder, if they had done their job properly.

There was ample evidence of police negligence in the investigation, with Sutcliffe a serious suspect as early as October 1977. Sutcliffe had been interviewed nine times over three years before his arrest. But the massive backlog of unprocessed actions allowed Sutcliffe's file to be hopelessly outdated and inadequate for police officers intending to interview him. Many times they did not know what had happened in previous interviews, or that any previous interviews had taken place. One policeman later revealed that on four occasions officers mistakenly believed it was the first time Sutcliffe had been questioned.

As Smith points out, Sutcliffe's boots left clear impressions at two crime scenes and were in his garage when he was interviewed. His alibis were given only by close relatives, including his wife.

> He had a previous conviction … as a result of being found in a woman's garden with a knife and a hammer – the weapons used by the Ripper. His appearance closely … matched photofit pictures created by two surviving victims. His various cars, which had left tyre prints at the scene of some … murders, were frequently logged by police in red light districts of Bradford, Leeds and Manchester [where the police expected to find the killer].[25]

Police failed to link the cases of his first three victims who survived with later murders, despite obvious similarities. Having decided they were dealing with a prostitute killer the police excluded genuine Sutcliffe victims and included a murder he didn't commit as a Ripper victim by reference to the 'character and habits' of the women involved, rather than relevant forensic evidence. Apparently the police expected the Ripper to be some kind of obvious monster and 'spent millions of pounds fruitlessly searching for an outsider when the culprit was just an ordinary bloke, a local man who shared their background and attitudes to a remarkable degree'.[26]

Sutcliffe was finally arrested and identified as the killer by chance, by police from a different force, not the West Yorkshire police who were conducting the investigation.

By refusing to allow any of these issues to be considered in court, the Law Lords completely 'begged the question' of police negligence, which was really, or should really have been, the central question at issue.[27]

The closely related fallacy of *complex question*

> is committed when a single question that is really two or more questions is asked and a single answer is then applied to both questions. Every complex question presumes the existence of a certain condition. When the respondent's answer is added to the complex question, an argument emerges that establishes the presumed condition.[28]

In law, complex questions are called *leading questions* (or double-barrelled questions) and rules of evidence exist to prohibit the use of such questions at certain times. Counsel is not permitted to 'lead' her or his own witness by using questions that suggest an answer or that assume unestablished facts in dispute: see, for example, *Evidence Act 1995* (NSW), s 37.[29]

On the other hand, as McNicol and Mortimer acknowledge, 'the witness must … confine themselves to the questions asked and presumably a well-thought out question will produce an answer favourable to that party's case'.[30] Most important, 'Unlike examination in chief, the general prohibitions on asking leading questions and attacking the witness's credit do not apply in cross examination'.[31]

It is easy to see the dangers here, with witnesses browbeaten and confused into incriminating themselves. Genuinely complex (double-barrelled) questions are objectionable – open to objection by the other side. But if the other side is asleep, judges can assume that they have consented to such questions.

In the past police interrogation has been notorious for use of complex questions. McConville et al set out four kinds of questions frequently encountered in police interrogations, through which 'facts' are created:

1 *Leading questions*. These seek to persuade the suspect to give a particular answer and to foreclose other possible answers. 'You intended to injure him, didn't you?', 'You set fire to the house deliberately, isn't that so?' are typical such questions.

2 *Statement questions*. These are statements which masquerade as questions; the suspect is confronted with a 'statement of fact' which the suspect is defied to contradict or invited to confirm. 'You took the money after displaying a knife and threatening the victim, didn't you?'

3 *Legal-closure questions*. These purport to invite the suspect to provide information but in reality force information into a legally significant category in the hope that the suspect will 'adopt' it. This may involve introducing some matter not previously mentioned or it may reshape what has been said so that it now 'fits' into an appropriate legal category. Thus, for example, where a suspect states that she took goods from a shelf in a shop and was apprehended outside the shop and had not paid but has not admitted that his was an intentionally dishonest act, the interrogation might continue 'So you stole the goods?' apparently re-stating what is already disclosed but in fact by supplying the conclusory concept 'stole', by-passing the need to establish the legal elements of 'intention to permanently deprive' and 'dishonestly' which are the basis of the offence of theft.

4 *Imperfect syllogistic questions*. The method involves persuading the suspect to accept the truth of a disputable or erroneous proposition by inducing the suspect to accept that it logically follows from acceptance of other (unarguable) propositions which have already been agreed. This will occur, for example, where the suspect agrees that an act was done which led to a particular consequence, and is persuaded to accept that guilt follows from these concessions (it often will not, without further findings relating, for example, to the suspects state of mind).[32]

The current requirement for videotaping interviews yielding admissions used in court only overcomes the problems of such complex questions if the court is fully aware of the issues involved here.

False dichotomy is another familiar fallacy of presumption.

> The fallacy of false dichotomy is committed when one premise of an argument is an 'either … or … ' [that is, disjunctive] statement that presents two alternatives as if they were jointly exhaustive [as if no third alternative were possible]. One of the alternatives is usually preferred by the arguer. When the arguer then proceeds to eliminate the undesirable alternative, the desirable one is left as the conclusion. Such an argument is clearly valid; but since the disjunctive premise is usually false, the argument is almost always unsound.[33]

Let us imagine that a bunch of white police have been caught on video engaged in an unprovoked attack upon an unresisting black motorist. The videotape creates major problems for the police defence counsel in a subsequent criminal trial. But false dichotomy comes to the rescue:

> Ladies and gentlemen of the jury, there is one simple question that you must address. Is it true that my clients are really the ruthless and inhuman criminal psychopaths they have been painted to be by the prosecutor, or are they

> merely ordinary decent people trying to do their best for the public in a difficult and dangerous situation?

There will probably be elements of the straw man here also. The prosecutor probably did not actually paint such a purely negative picture. But the crucial point is to get the jury to accept that these are the only two possibilities. It has to be one or the other. So if they reject the first, as they are likely to do if it is obviously exaggerated, false and unfair, they have to accept the second.

A basically similar argument was presented by defence counsel in the criminal trial following the police beating of Rodney King in Los Angeles. Whether or to what extent it influenced the jury is unclear because of the lack of any requirement for written justifications or explanations of jury verdicts.

Arguably, while it is important to shield juries from external pressures, the inscrutability of jury decisions is a major anomaly in a supposedly rational common law system. If jurors have made some serious error in reasoning then it should be open to challenge and a possible basis for appeal.

Discussion topics

1. What are fallacies of weak analogy? Explain with reference to the *Hill* case.
2. What are fallacies of hasty generalisation, causality and presumption? Include reference to police interrogation.
3. Seek out and analyse inductive fallacies and fallacies of causality and presumption in textbooks, newspapers and judgments.

Additional resources

R. Graycar and J. Morgan, *The Hidden Gender of Law*, Federation Press, Sydney, 2003 (on police negligence cases in different jurisdictions).

M. McConville et al, 'The Case for the Prosecution' (1991), in D. Brown et al, *Criminal Laws*, 2nd edition, Federation Press, Sydney, 1996, pp 204–7.

For interesting parallels between police and advocates' immunity, see *Swinfen v Lord Chelmsford* (1860) 5 H&N 890; *Rondel v Worsley*

(1969) 1 AC 191; *Giannarelli v Wraith* (1988) 165 CLR 543: FC 88/047; *Boland v Yates Property Corporation Pty Ltd* (1999) HCA 64 (particularly Callinan J's observations on *Giannarelli*); *Arthur JS Hall & Co v Simons* (2000) UKHL 38 (20 July 2000); *D'Orta-Ekenaike v Victorian Legal Aid* (1997) 2 SLR 729 (comparing the comments of Kirby J with those of Callinan J).

Chapter 5

SCIENCE AND STATISTICS*

Science and law

It is more difficult to provide any sort of simple and concise definition of science parallel to the definition of logic provided in chapter 1. Clearly, this chapter cannot consider all the different disciplines, ideas, theories and research practices of modern natural and social sciences. Its aim is the more modest one of focusing upon key elements of *scientific method*. These include methods of producing and reliably testing scientific theories (or producing reliable scientific knowledge) and applying such theories in explaining otherwise inexplicable facts and observations.

There are problems here too, insofar as many people have argued that there is no such thing as scientific method, rather a host of different sorts of methods employed by different sciences, with no single set of ideas or practices common to them all. But we will be focusing upon what are widely argued to be central structural features of methods employed across a range of different scientific disciplines. We have, in fact, already taken some significant steps in that direction in earlier chapters, insofar as *deductive and inductive logical reasoning* processes are integral to all scientific method. The next step is to consider the precise role of such logical reasoning in scientific experiment, scientific theory testing, and scientific explanation.

Here, in particular, we will consider the formulation, testing and explanatory application of what have been called *theoretical hypotheses* (such as Newton's theory of gravity and Darwin's theory of evolution) and *causal hypotheses* (such as the claim that high

intensity magnetic fields cause cancer, or free trade agreements increase the gross domestic product (GDP) of all parties). We will see why a basic knowledge of *statistics* is necessary to understand causal hypotheses and why it is increasingly important for lawyers to be able to make sense of such hypotheses.

We will see that it is not necessary to understand the specific details of individual theories in order to able to critically assess and respond to scientific evidence of all kinds by reference to basic ideas of scientific method and scientific reasoning. And where it is necessary to gain a deeper understanding of such specific details, previous acquaintance with such basic issues of method and reasoning can be extremely helpful.

The previous two chapters identified a number of areas where science impinges upon the theory and practice of law. Four of these areas will be considered in greater depth in this chapter and subsequent chapters. They are:

1 the influence of ideas of scientific method upon legal methods;
2 problems of scientific evidence in the courtroom;
3 the role of scientific evidence in establishing causation, in civil and criminal law;
4 the role of social scientific research in law reform.

Many crucial issues of the relation between science and law are not covered here, most notably issues of effective control and legal ownership of scientific knowledge and tools of scientific research.

Science and anti-science

Some people say that science, like religion, rests upon certain articles of faith or fundamental (unquestioned) assumptions, to the effect that the world has a determinate structure that is discoverable by, and comprehensible to, human beings, using appropriate methods. The methods that are most effective in this regard are the methods of science, involving the use of human faculties of perception, introspection, memory and reason, in ways that are, in theory, open to everyone. In fact, the progress of modern science provides 'rational' empirical support for such claims, making them more than mere 'assumptions' or articles of faith.

Many people today associate science with a mechanistic, materialistic and atomistic world view, which encourages and sustains the

manipulation and destruction of the natural world in pursuit of private profit. They argue for the adoption of more organic, holistic and ecologically responsible perspectives.

But as Theodore Schick and Lewis Vaughn point out, while particular world views are dominant in particular scientific communities at particular times, 'it would be a mistake to identify science with any particular worldview'.[1] For science, or at least the part of science that concerns us here, is nothing more than a 'method of discerning the truth', along with the body of, at least provisionally, established truths produced by application of that method. It is, in fact, the best or only method for discovering the deeper structures or mechanisms that lie behind the world of immediate, day-to-day experience, which generate and shape that world, and reference to which therefore explains such experience.

> Those who believe that we should adopt a more organic and holistic worldview do so on the grounds that it offers a more accurate description of reality than does a mechanistic and atomistic one. That may well be the case, but the only way to find out is to determine whether there is any evidence to that effect, and the best way to make such a determination is to use the scientific method.[2]

It may be the case that this powerful method is currently employed primarily in the service of private profit. Current political-economic arrangements may render some results a source of greater harm than good. But both the method and the results can, and should, be used for the benefit of all people, of society and of the natural world.

Naive inductivism

For a long time scientific method was assumed, by many, to be a simple four-stage process. First of all we observe the world, and accumulate a mass of singular observation statements such as:

> At time A on date B, the planet Venus appeared in position X, Y, in the sky above Sydney.
>
> This A-centimetre long bar of iron was seen to increase its length by B centimetres when heated from Y to Z degrees Celsius.

Providing certain conditions are satisfied, we then generalise from a finite list of singular observation statements to a *universal hypothesis*. From a list of observation statements referring to the expansion of specific heated metals, for example, we generalise to the universal hypothesis that

All metals expand when heated.

The conditions that must be satisfied for legitimate generalisation include the following:

The number of observation statements forming the basis of a generalisation must be large.

The observations must be repeated under a wide variety of conditions.

No accepted observation statement should conflict with the derived universal law.[3]

We can then deduce specific things that must be true if our hypothesis is true:

This newly discovered metal will expand when heated.

Railway lines will buckle on hot days if inadequate space is left between them.

And we can verify, or refute, the hypothesis by experimental testing of such deduced implications. If it survives such testing it comes to be regarded as a universal 'law' of nature.

Such universal generalisations are valuable resources to the extent that they enable us to make reliable predictions about the future and provide reliable explanations of current observations – the kind of explanations considered in chapter 1. The rail buckled because the temperature exceeded X degrees. Similarly laid tracks will buckle in similar fashion when the temperature exceeds X degrees in the future.

We suggested in chapter 1 that such a model of inductive generalisation probably influenced judicial thinking in the later 19th and earlier 20th centuries. Because it appeared to have been so successful in the natural sciences, it was thought that it should therefore be adapted and applied in law. Hence the search for universal principles underlying specific judgments in particular areas, culminating in the neighbour principle and the *Fletcher v Rylands* principle. Such general principles could then be applied, not in prediction and explanation, but in the simplification and ethical rationalisation of legal reasoning processes, now reduced to a simple matter of deduction (of application of *modus ponens*), rather than complex and uncertain reasoning by analogy. Instead of having to balance lots of difficult similarities and dissimilarities (in analogical reasoning) it now becomes a matter of checking off a list of necessary conditions.

As we will see below, there are good grounds for regarding analogical reasoning, rather than simple inductive generalisation, as a major

mechanism of theory construction in science. But we also saw in chapter 1 that inductive generalisation and analogy are closely related. And we should not be too quick to completely reject the significance of inductive generalisation in science, as some have tried to do. Newton's 'discovery' of universal gravitation seems to have involved analogies between a falling apple, a cannon ball and the moon, but so did it involve a sweeping generalisation drawing out essential and by no means obvious features common to falling apples, flying cannon balls, the orbiting moon and a host of other earthly and celestial motions.

Lord Atkin's derivation of the neighbour principle looks more like Newton's derivation of the gravity principle through abstraction of deep structural principle from diverse cases rather than the simplistic 'all metals expand on heating' type of inductive generalisation. He has seen beyond the concrete details of dogs biting neighbours, surgeons chopping off the wrong limbs and manufacturers selling drinks full of snails to abstract structural-ethical features shared by all such disparate fact situations.

Scientific theories

There are still many problems with the simple inductive/deductive model as a paradigm of scientific method, highlighted by philosophers of science since the mid-20th century. It is frequently argued that scientific enquiry typically starts with some sort of *question or problem*, considering why something occurred (Why did the dinosaurs become extinct?), why some thing has the particular properties it does have (Why do metals expand when heated? Why do the continents fit together like a jigsaw puzzle?), or how particular things are related (for example, mammals and reptiles, electricity and magnetism), rather than engaging in undirected observation. Observations are needed to recognise that such problems exist, but it is the attempt to solve the problem or answer the question that stimulates the production of (explanatory) hypotheses of some kind, along with subsequent testing of such hypotheses.

There are also serious problems with the naive inductivist assumption that theories are straightforwardly verified or refuted by new observations, consistent with, or contradictory to, the theory. As Schick and Vaughn note, 'when new hypotheses are first proposed, there is often a good deal of evidence against them'.[4] This is because

both the theory itself and the background of other theory upon which it depends have not yet been adequately developed to address such counter-evidence.

Other serious problems revolve around the simplistic empiricist *epistemology*, or *theory of knowledge*, built into the model, which suggests that scientific knowledge of the world can be gained through purely passive observation and generalisation without the need for any kind of directed and active intervention in theory or in material practice.

As noted in chapter 1, science involves the search for general causal principles, in terms that we can use to explain and predict the world as perceived. We have already briefly considered the nature of causes and causal explanation. But to better understand science and scientific method we need to look at bit more deeply into what is involved.

Both science and everyday observation tell us that the world is made up of relatively enduring and qualitatively distinguishable types of things: gases, liquids and solids; bacteria and fungi; plants and animals; and stars, galaxies and clusters of galaxies. They also tell us that that each different 'natural kind' of thing possesses particular 'causal powers' or capacities or tendencies to exert sorts of force, able to influence other things – as well as themselves – in characteristic ways. And they possess such powers by virtue of particular internal and external structural relations that make them the kind of things they are.

Material bodies exert powers of mutual gravitational attraction (directly proportional to the product of their masses and inversely proportional to the distance between their centres of mass) or powers to respond to the local curvature of space. Once in motion, they move themselves along with a constant velocity in a straight line unless acted upon by an external force, and resist the action of such a force in proportion to their mass. They respond to such an external force by accelerating in proportion to, and in the direction of, that force. Liquids have the power to flow downhill, dissolve various sorts of solid, freeze at low temperatures and turn to gases at higher temperatures. Chemical elements can combine with other elements in particular proportions to create new sorts of substances with different properties from the elements in question. Plants and animals can grow and develop and reproduce. Animals and humans can move, and sense and think. Such powers are released or blocked depending upon the particular configuration of types of things involved. Certain conditions are 'necessary' for release of the powers in question.

Some such intrinsic powers do reveal themselves to the sort of passive observation considered in the simple inductive/deductive model (of scientific method) insofar as they are triggered into action without the need for human intervention to produce directly observable effects. Metal rails expand as a result of heating by the sun and by friction with train wheels, iron turns to rust through interaction with air and water, animals mate and produce offspring. But in many cases, to gain a deeper understanding of the true nature of the powers in question, substantial active intervention by people is required. This is nicely illustrated by the classical physical ideas of gravity and inertia.

Everyday observations tell us that unsupported bodies generally do fall, though not if they are very light or the wind is very strong. But to observe – and measure – the effects of gravity uninfluenced by friction of earth and atmosphere requires the isolation of a gravitational system from other such – predominantly electromagnetic – influences. Such isolation can, initially, be achieved in the imagination, in a 'thought experiment'. And particular 'real world' experiments can help with such mental extrapolations. Galileo achieved this (viz-a-viz gravity) by extrapolation from observations of bodies falling in media of decreasing density and rolling down slopes of increasing inclination. But beyond this point, countervailing forces must be effectively excluded or balanced against each other to approach a 'pure' gravitational interaction, as when a feather is dropped in a vacuum chamber or on the surface of the moon or when we look out into the vacuum of space.

The same is true of inertia. Everyday observation tells us that heavy objects are more difficult to redirect once they get moving at a reasonable pace. But on the earth we never directly perceive the tendency of masses to continue in uniform motion in a straight line so long as no forces act upon them, because forces always do act upon them. Again we can extrapolate from movement over surfaces of decreasing friction. And again, countervailing forces must be excluded or balanced to get close actually seeing inertial motion. An orbiting spacecraft is still affected by the gravitational force of the earth. But in the frame of reference of objects within such a craft, this gravitational force is counterbalanced by an equal and opposite centrifugal force, leaving no net force upon such objects and rendering them weightless. So they do, indeed, exhibit uniform inertial motion.

In many cases we discover the causal powers of things only by creating special 'artificial' situations that trigger or release powers which

normally lie dormant by virtue of the ubiquitous presence of countervailing or blocking forces in our immediate environment. We release such powers by blocking off the blocking forces. The technologies of internal combustion engines and atomic reactors, for example, are based upon the creation of special environments that release the (generally dormant) heat energy-producing powers of petrol and uranium (dormant on the surface of the earth at least).

In many other cases, we observe correlations (or concomitant variations) of various kinds, or situations of contiguity (where A-type events are regularly found together with B-type events) without knowing whether, or what sort of, a causal relationship is involved. More smokers than non-smokers are observed to develop lung cancer. Lung cancer rates appear to be directly proportional to quantities of cigarettes smoked, richer people on average live significantly longer than poorer people, inflation seems to rise in inverse relation to unemployment. Chinese herbs appear to produce good results in some cases. But does smoking cause cancer? Are Chinese herbs effective medicines? Again, the answers to these sorts of questions cannot generally be found by simply – passively – accumulating more observations of the same kind. Active intervention is required to decide such questions, through application of appropriate procedures of controlled experimentation, or complex analysis of existing data from a variety of different sources.

In all of these cases we are indeed concerned with the production of a particular theory or hypothesis on the basis of particular sorts of observations, as suggested in the simple inductivist model. The theories in question, dealing with entities and causal powers and tendencies of such entities that are – more or less – directly observable, have been called *Type 1 theories* (to distinguish them from *Type 2 theories*, dealt with in chapter 7).[5] They include theories about the biological, chemical, gravitational, electrical and magnetic properties of observable things (moving electric charge gives rise to magnetism, changing magnetism gives rise to electricity), about the pressure, temperature and volume of gases ($pV = nRT$), about the physiology, development, behaviour and circumstances of animals and people, including the role of particular substances or experiences or life circumstances in causing or curing illness (Chinese herbs cure illness, smoking causes lung cancer) and about the interaction of such observable and measurable social variables as GDP, inflation, inequality, morbidity, mortality and violent crime.

Rather than further passive observation, the subsequent testing of such theories typically involves a range of different sorts of experimental investigation, actively seeking relevant data through surveying and sampling, and subjecting selected phenomena to isolation and controlled manipulation.[6]

Testing Type 1 theories: proportions

Here we concentrate upon a particular subset of Type 1 theories that raises special problems (qualitatively different from those considered in relation to Type 2 theories,) by virtue of the necessary involvement of statistics. Many law students and lawyers recoil in horror at the prospect of mathematical complexities. But the sorts of hypotheses considered here, *simple statistical hypotheses*, *correlations* and *causal hypotheses*, are, as Giere points out, 'the stock in trade of psychology and the social sciences, and of science related fields such as bio-medicine, public health and education'.[7] They are, therefore, precisely the kind of ideas with which lawyers are increasingly likely to come in contact, both in relation to issues of factually based law reform, and expert testimony in the court room. And, luckily, the basic principles of theory testing in this area can be grasped with minimum statistical detail.

Considering first simple statistical hypotheses, we see that all such hypotheses identify 'a population, a property that each member of the population may or may not exhibit, and a percentage'. As Giere says:

> Each of these statements could therefore be expressed in a statement with the following structure:
>
> X per cent of (Population) are (Property).
>
> The percentage tells us the relative number or *proportion* of the members of the population that exhibit the property in question.[8]

Here are some examples:

- 12% of sentenced prisoners in gaol in Australia on 30 June 2000 had been convicted of sexual offences.[9]
- 4% of male prisoners serving a sentence of 12 months or more in the NSW prison system in 1973 had completed the HSC.[10]
- 80% of female prisoners in NSW prisons in 1996 had been in at least one violent relationship.[11]

Simple statistical hypotheses can be seen as answers to specific sorts of questions:

- How many (what proportion of) apples in this consignment are rotten?
- How many (what proportion of) Australians are now HIV-positive?
- How many (what proportion of) Australian teenagers are cigarette smokers?
- Are more people in prison for violent than non-violent offences (that is, more than 50 per cent)?
- How many (what percentage of) men in Australian prisons are raped each day?

There are three (fundamental) scientific, issues that arise in relation to these sorts of questions:

1 *Providing appropriate – operational – definitions for the properties in question, as observable and measurable properties*. What counts as a rotten apple? How precisely do we distinguish the rotten from the non-rotten? What counts as a violent relationship? How many cigarettes make you a smoker? What if you don't inhale?
2 *Identifying and accessing the relevant population*. Who counts as a teenager? Who counts as a male prisoner?
3 *Selecting an appropriate sample*. Do we test the whole population? If not, how big a sample do we choose and how do we select the sample?

Special problems are involved in questioning people in order to illicit the relevant information, questioning them about their beliefs, habits, past experience, present emotional/physical condition or future plans (as opposed to counting apples). People may misremember facts, they may be mistaken about their own future actions or unwilling to tell the interviewer the truth, for one reason or another. Special efforts may be required to try to avoid or overcome these problems.

Sometimes we can check a whole population. For example, beginning in 1982, a national prison census has been conducted each year in Australia on 30 June. The census was administered by the Australian Bureau of Criminology until 1993 and thereafter by the Australian Bureau of Statistics.[12]

Our first example of a simple statistical hypothesis came from the year 2000 census, which involved all prisoners in Australia. This sort of thing is informative but also costly and difficult to organise. Very often we cannot check the whole population of interest, and must instead take and check *a sample*. On this basis, we make an *inductive generalisation* about the whole population.

For a statistical induction to be successful, the sample must adequately represent the population. It is not enough that the sampled individuals be of the same category (apples, prisoners, teenagers); they must also have the same mix of sub-categories (or 'secondary properties' – male, female, old, young) which may be relevant in determining the 'primary properties' of interest to us. Such a mix can best be achieved through ensuring that our selection is made randomly, with each member of the population (who might exhibit the property of interest) having an equal chance of being selected (and no correlation between the outcome of one selection and another).

How we achieve such random sampling depends upon the particular population we are sampling. It might involve an appropriate mixing of the entities in question or it might involve assigning numbers to each of them and using random number tables to select some. The latter sort of approach was adopted in two other prison surveys that yielded data referred to above. The NSW Bureau of Crime Statistics survey of 1973 involved a random sample of 1000 male prisoners serving a sentence of 12 months or more. The 1996 NSW Corrections Health Service Inmate Health Survey involved a randomly selected sample of 789 prisoners from NSW gaols.

P1 A random sample of 789 prisoners was collected from NSW gaols.

P2 One-quarter of the women prisoners sampled reported having used heroin in prison.

C So probably close to a quarter of women prisoners in NSW gaols have used heroin while in gaol.

Even with a random sample, probability theory tells us that the sample probably will not directly or precisely reflect the distribution of the property of interest in the population as a whole. Suppose that 50% of a population of people have at some time been victims of property crime. So the probability of being such a victim in the population is said to be 0.50.[13] But the chance that our sample will precisely mirror the proportions in the population is actually rather slim. One sample of 100 might yield 55 victims, another 47. Even if we are

careful to choose a representative sample, this in no way guarantees the proportion of such victims will be exactly 0.50.

Such 'sampling error' cannot be eliminated. But it can be reduced by taking sufficiently large samples. Furthermore, a major statistical result, called the *central limit theorem,* says that the (probability) distribution of all possible sample proportions (or sample means) – indicating the different probabilities of each possible proportion – is normal, as long as the sample sizes are large enough, that is, over about 30.

In the bell-shaped normal distribution, the mean or average value is the central high point of the curve, which also corresponds with the median or central value. In a normal distribution of all possible sample proportions (47/53, 48/52, 49/51, 50/50, 51/49, etc) this mean value corresponds to the actual population proportion (or mean). Furthermore, the variability of the sample proportions, that is, the spread around the mean value, decreases as sample sizes increase.

Here, we need to briefly consider the key statistical concept of *standard deviation*, as a kind of average of individual deviations from the mean of a distribution, or measure of how widely dispersed the values are (with big standard deviation = big dispersion). To calculate the standard deviation of a variable in a population or a sample, we add the squared deviations from the population or sample mean, divide by the size of the population or sample and then take the square root of the result.

The so-called *empirical rule* states that in a normal distribution, approximately 68% of values lie within one standard deviation of the mean, or between the mean minus one standard deviation and the mean plus one standard deviation. Approximately 95% of values lie within two standard deviations (either side) of the mean, and 99.7% within three standard deviations (either side) of the mean.

Population values deviate from each other due to natural phenomena. Sample values vary because of the errors that occur because the whole population has not been considered. The variability of the latter is therefore referred to as *standard error* (SE) rather than standard deviation. But the empirical rule applies also in relation to such SE in sample distributions. SEs are calculated by reference to the value of the population variable in question (in this case proportion – p = 0.50 victims of property crime) and the size of the samples involved, n (in this case 100). Thus, SE of the distribution of sample proportions is square root of [p(1 – p)/n]. We see that the SE decreases as n, the sample size, increases.

Because the sampling distribution of sample proportions (or means) is normal, we can use the empirical rule to determine how much a given sample proportion is expected to vary from the true population proportion. Applied to the normal distribution of all sample proportions (for any given size of sample over 30) the empirical rule says that we can expect about 68% of sample proportions to lie within one SE of the population proportion, about 95% to lie within two SEs, and about 99.7% to lie within three SEs of the population proportion.

Generally, we want to discover the population proportion by reference to a single sample proportion. We know that 95% of all sample proportions lie within two SEs of the population proportion. So if our estimate is actually a range including the sample proportion plus or minus two SEs, our estimate should be correct about 95% of the time. The number of SEs added or subtracted is called the *margin of error* (ME). Thus, the ME is supposed to measure the maximum amount by which the sample proportion (or other statistic) is expected to differ from that of the population. By combining our sample proportion, as estimate of the population proportion, with the margin or error, we come up with a confidence interval. One SE either side of our sample estimate gives a confidence interval of 68%: we can be 68% confident that the actual population proportion is within an interval of one SE either side of the sample proportion. Generally, statisticians go for a 95% confidence level, which means an interval two SEs either side of the sample proportion.

The problem is that SE depends upon population data that will generally not be available: in this case, the actual population proportion. This is what we are probably trying to discover, on the basis of sample data. As it happens, on the assumption that our sample is, indeed, representative of the population, we can estimate the SE proportion by reference to the observed proportion (or 'frequency') in our single random sample. Estimated SE proportion = square root [(f)(1 – f)/n], where f = observed sample frequency and n = sample size.

For example:

For a sample size of 200, and observed frequency of 0.4:

SE proportion is the square root of [0.4 x 0.6]/200 = 0.035;

0.035 = 1 SE = confidence level of 67% = 80 plus or minus 7;

0.07 = 2 SEs = confidence level of 95% = 80 plus or minus 14.

By assuming the sample frequency to be 0.50, we can produce tables of approximate MEs at the 95% confidence level for different sample sizes. Such tables give over-simplified results, but they can be extremely handy for purposes of rough calculation, and speedy assessment of relevant data.

Sample size	Margin of error
25	+/– 0.25
100	+/– 0.10
500	+/– 0.05
2000	+/– 0.02
10 000	+/– 0.01[14]

Larger sample sizes give better information. A larger sample with the same ME gives a higher confidence level. A smaller sample with the same confidence level gives a larger ME. Higher confidence requires better information.

Correlations (1)

So far, we have considered only single variable properties of members of particular populations, and how to generalise from the observed properties of a sample to a whole population. The next step is to consider the relationships between two or more variable properties of individuals. Consideration of the strength of relationships between variable properties brings us into the study of correlations.

Correlations are concerned with the strength of the relationships between the values of two variables. A high positive correlation means close to direct proportionality: one value increases at the same rate as the other. A high negative correlation means close to inverse proportionality: one value decreases as the other increases. The nature and extent of such associations are measured in terms of what are called *correlation coefficients* (CC). A perfect positive correlation corresponds to a CC of plus 1, a perfect negative correlation to a CC of minus 1.

Although a relationship of correlation is not the same thing as a relation of causation, correlations are generally indicative of causal relationships of some kind. Depending upon how we define causation, correlation could be a necessary feature of a causal relation, though not, in itself, sufficient to demonstrate causation.

To establish the existence of a correlation is often to take a first step towards establishing a causal relationship, though the correlation itself does not necessarily indicate the nature of the causal relationship involved. As Philips points out, when variable X and variable Y are substantially correlated, it might be that X is causing Y; it might be that Y is causing X; it might be that both are caused by a third variable or set of variables or that each is a single aspect of a 'multifaceted but indivisible whole', as with the similar speeds of the two front wheels of a car.[15]

An example here would be a sample of circles of different radii. As Rowntree notes, 'the circumference increases in length when the radius increases. Big values of one variable are associated with big values of the other, and small with small'.[16] If we plot circumference in centimetres on the y axis against radius in centimetres on the x axis, we get a straight line (rising from left to right), indicating a very strong correlation. But it is misleading to say that either variable causes the other. Rather, the formula $C = 2\pi r$ describes a fundamental structural feature of all circles.

Another example involves the variables of acceleration and force in relation to a population of objects of similar mass (or indeed, the same object subjected to different degrees of force). Here, following appropriate experimental procedures to generate the results in question, we plot force (f), increasing from one to five (arbitrary) units, on the x axis, against acceleration (A), in m/sec^2, on the y axis. And again, we get a straight line rising from the left to right that would, in this case, pass through the origin.

Acc (m/sec^2)	Force
0.20	1.0
0.40	2.0
0.60	3.0
0.80	4.0
1.00	5.0[17]

This suggests a direct relationship between force and acceleration. A is proportional to f, when the mass of the object is held constant.

Next we consider objects of differing mass, with the force kept constant (at one unit). We again plot acceleration on the y axis, with mass on the x axis:

Acc (m/sec^2)	Mass
0.8	1
0.4	2
0.210	3
0.200	4
0.160	5[18]

This time, we get a concave curve falling from right to left. This suggests an inverse relationship between mass and acceleration (with force held constant). It seems that A is proportional to 1/m.

The next stage is to plot f/A (on the y axis) against m (on the x axis). Again we get a straight line rising from left to right (intersecting the origin) and suggesting direct proportionality – f/A proportional to m. In fact, f = kma, and we can choose units that make k = 1, to give Newton's second law, f = mA. And in this case, the correlation does point to a causal relationship, with applied force causing acceleration (and mass resisting such acceleration).

Another straightforward example involves increasing volume (V) of a gas, plotted on the y axis, against the pressure (p) exerted by that gas, plotted on the x axis, with the temperature and amount of gas particles held constant. Here again appropriate experimentation yields results that give a concave curve falling from left to right (basically similar to the mass/acceleration graph) suggesting that V is proportional to 1/p. And this is confirmed when we plot A against 1/p to get a straight line graph. This is known as Boyle's Law, and is again indicative of a causal relationship, in this case involving the impacts of gas particles on the walls of the vessel containing the gas.

Finally, to see that correlations are significant in the social, as well as the natural sciences, we consider the quantitative studies of Lars Schoultz, Michael Stohl, David Carleton and Steven Johnson,[19] examining the relationship between US foreign aid and observance or non-observance of human rights by receiving governments between 1962 and 1983. As Gareau points out:

> The human rights [violations] tested for in the first study [covering the years 1962 to 1977] were torture and other forms of cruel, inhuman, and degrading treatment, including prolonged detention without trial … The problem of determining the level of observance or non-observance of human rights was left in this study to the composite judgment of 38 experts from Western non-communist countries who had

> published widely on the subject or who had occupied key positions in non-governmental human rights organisations. The experts' mean evaluation of human rights violations for the 23 Latin American countries serve as the dependent variable … The independent variable was aid from Washington … The correlations between the level of human rights violations and aid in the first study were found to be 'uniformly positive', indicating that aid had tended to flow disproportionately to Latin American governments that torture their citizens. In addition, the correlations are relatively strong … The seven countries that received little aid … maintained a relatively high level of support for human rights, while the 16 that received comparatively large amounts of aid had a low level of respect for these rights.[20]

The second study found a direct relationship between foreign assistance and human rights violations during the Nixon and Ford administrations. 'The more violations, the more aid received.'[21] And the third study, covering 59 countries over the period 1978–83, found 'that no matter what scale was used for human rights violations no significant negative correlations with aid from Washington could be found for any year tested'.[22] In some cases, 'the more violations that occurred, the more aid was forthcoming. This was found to be the case both for economic and for military aid'.[23]

Correlations (2)

Here we concentrate upon a more restricted version of the idea of correlation that allows us to carry over the ideas of sampling and margin and error considered earlier. In a population whose members have properties P (and not P) and Q (and not Q), P is positively correlated with Q when the proportion of Q's amongst the P's is greater than the proportion of Q's amongst the non-P's.

As far as sampling is concerned, exactly the same principles apply as with simple proportions. We need to aim for a random sample, and to take account of appropriate margins of error. Sticking with our example of crime victims, we might consider two variables: victimisation status in respect of crimes of violence and employment status. We assume that each has two possible values: victim or non-victim, and employed or unemployed.

When the percentage of victims is the same among both employed and unemployed people, we can say that the two variables – victimisation and employment status – are not correlated. If the percentage of victims amongst the unemployed is greater than the percentage

amongst the employed, we can say that for this population, being a victim of violent crime is positively correlated with being unemployed.

Here again, Giere provides a concise analysis of precisely what is involved. Let's say, in a sample of 600 people from a particular region 500 turn out to be unemployed and 100 employed. Amongst the unemployed, 325 turn out to be crime victims. In other words, f(V/U) = 65% (frequency of victims amongst the unemployed). Amongst the employed people in the sample, 35 out of 100 are victims, so f(V/not U) = 35% (frequency of victims amongst the employed).

There is a clear difference in the sample between these two values, f(V/U) and f(V/not U). The question is whether this difference in sample frequencies constitutes good evidence for the existence of a correlation in the population. To find out it is necessary to construct the relevant interval estimates (the relevant ME's)

By our rules of thumb (established earlier in relation to distributions), ME for a sample of 500 (unemployed people) = 0.05. So we estimate P(V/U) = f(V/U) +/– ME = (65 +/– 5)%. Similarly for n = 100, ME = 0.10. So we estimate P(V/not U) = f(V/not U) = (35 +/– 10)%.

If we check the lowest possible value for P(V/U), that is, 60%, against the highest possible value for P(V/not U), that is, 45%, we see that there is no overlap between these two distributions (or sets of values). That is good evidence for concluding that P(V/U) really is larger than P (V/not U) in the population. Being a victim really is positively correlated with being unemployed. You are more likely to be a victim of violence if you are unemployed, and more likely to be unemployed if you are a victim of violent crime in this imaginary population.

If we knew precisely how many unemployed people were victims, and how many were not, we could use the difference between these two figures as a measure of the strength of the correlation between being unemployed and being a victim. For example, if 75% of unemployed people were victims and 25% of employed people were victims, then we could measure the strength of the correlation as 0.75 – 0.25 = 0.50, which is a fairly strong correlation. Measured in this fashion, correlations would range from 1.00, where all the unemployed were victims and none of the employed were, through to 0.00 for non-correlation, and to –1.00, where none of the unemployed are victims and all the employed are.

Where we do not know the exact proportions, but only the end points of 95 per cent confidence intervals, then we can use the end points of the two intervals to determine both the maximum and the

minimum difference allowed by the two intervals. The maximum allowed difference is between the top of the higher interval and the bottom of the lower interval. The minimum allowed difference is between the bottom of the higher interval and the top of the lower interval.

In the example above, the maximum allowed difference in proportions is 0.70 – 0.25 = 0.45. The minimum allowed difference is 0.60 – 0.45 = 0.15. So the estimated strength of the correlation is the interval (0.15, 0.45). This does not include zero, and therefore non-correlation is excluded.

As noted earlier, this is a rather restricted idea of correlation. A broader version of correlation is a relation between sets of variables, for example, numbers of police per 10 000 population in different countries, compared to violent crimes per 10 000 population in these countries. There are a number of statistical procedures that can be applied to calculate the extent of the correlations in question, with direct proportionality measured as plus one, inverse proportionality as minus one and no apparent relationship as zero. Perfect positive correlations correspond to straight lines rising left to right when the variables are plotted on x and y axes, falling left to right for perfect negative correlation.

Discussion topics

1. What is science? How is it related to law? Include reference to anti-science, junk science, naive inductivism and scientific theory.
2. What are Type 1 theories? Include reference to proportions, correlations and causation.
3. Find and discuss some references to theories you now know to be Type 1 theories from textbooks, newspapers and cases.
4. What is statistics? Include reference to random sampling, normal distributions, standard deviation, margin of error and significance testing.

Chapter 6

CAUSATION AND THE PRECAUTIONARY PRINCIPLE

Causal hypotheses

Quite a lot has already been said about causation. Scientific researchers generally act as though everything has a cause, and they aim to find the causes in their chosen area of research. At the same time, they distinguish between strictly deterministic causes, where A is sufficient for B, and stochastic causes, where the occurrence of A merely increases the likelihood of the occurrence of B.

As noted earlier, we typically identify a range of different conditions as individually necessary and jointly sufficient to bring about particular effects. An individual's ankle pain is the result of overexertion in the tennis game just completed, the hardness of the court, the length of the game, a genetic weakness in the ankle and a fall.

To take another example, it seems that cancers typically develop only through a number of different mutations in an individual's genome. Some such mutations might be inherited, but others will result from exposure to a range of different environmental carcinogens. Only with all of the sequence established will the cell become cancerous, and only an immune system weakened by other environmental stresses will allow the cancer cells to proliferate.

It is causation in this sense, in which every event is recognised as 'the result of many conditions that are jointly sufficient to produce it',[1] that is the typical object of legal consideration. As Fleming states:

> in legal enquiries it does not matter if we are unable to identify all, or even most, of the individual elements which constitute the complex set of conditions jointly sufficient to produce the given consequence. The

> reason is that we are usually interested only to investigate whether one, two or perhaps three specific conditions [for example, identified acts or omissions by the defendant or other participants in the accident] were causally relevant … Whether a particular condition qualifies as a causally relevant factor will depend on whether it was necessary to complete a set of conditions jointly sufficient to account for the given occurrence.[2]

We also recognise the existence of long chains of causation, as, for example, where the better tennis court was closed for repairs, so the tennis player had to go to a different (harder) court, where they had to make greater than usual efforts to compete with a better player, leading to a fall, leading to greater stress on the leg, etc.

Such chains of causation are frequently objects of legal consideration. As Fleming points out:

> the defendant's default must be accounted a '*proximate*' *cause* of the [plaintiff's] harm, the consequence must not be too 'remote' [if legal liability is to be recognised]. And what is meant by 'proximate' [is] that because of convenience, of public policy, or a rough sense of justice, the law arbitrarily declines to trace a series of events beyond a certain point.[3]

We have also seen how we tend to pick out certain factors (from clusters of necessary conditions) as 'the causes' of particular events, typically changes shortly before the events in question, as opposed to standing conditions, which are constant for longer periods. Assumptions and beliefs of people looking for explanations are clearly relevant here. We know there is oxygen around, we are looking for some other reason for the fire. But the unknown conditions of interest are not necessarily the most recent changes. We might want to know why the tennis player was playing upon a different court from their usual one.

Reproductive disorders in seal species in the Baltic Sea in the 1970s might turn out to be the results of the production and sale of longlasting toxins (for example, DDT and PCBs) in US chemical plants years or decades earlier. The death of a particular person from mesothelioma cancer in one country could be traced back to a brief workplace exposure to asbestos decades earlier in another country. A violent act of abuse by an adult today might be, in part, a consequence of abuse that they suffered in childhood.

Causes can also be connected in positive and negative feedback loops, the former leading to accelerated change (as where, as Robert Martin suggests, plant death causes soil erosion, which causes more

plant death and more soil erosion), the latter to stable equilibria (as where a nation's trade deficit leads to devaluation of its currency through sales in international money markets, which in turn leads to cheaper exports and the removal of the deficit – supposedly). Positive feedbacks are particularly implicated in global warming, with melting ice reducing the reflection of sunlight from the ocean and land surface. This leads to further heating of the water and land and melting of the ice, melting permafrost releasing methane, which accelerates global warming, leading to further melting. Further melting leads to increased CO^2 accelerating plant growth in peat bogs, which in turn leads to break down of the peat and further release of CO^2.

As noted above, correlation is typically a sign of causation. This is clear in the example of force, mass and acceleration in the previous chapter. There we considered experimental 'application' of different forces to a fixed mass to produce different accelerations.

But there are many possible complications here, as nicely outlined by philosopher Robert Martin. We might find a positive correlation between measles spots and bodily aches and pains, not because either is the cause of the other, but because both are the effects of the progress of the infection. We might find a negative correlation between sunburn and mosquito bites, not because mosquitos don't like burned skin, but because hot dry summer weather reduces the water available for mosquito development and causes sunburn. We might find a negative correlation between eating lots of vitamin C and catching colds, not because vitamin C protects against colds but because the same people who take vitamin C also tend to do other things that reduce their chance of catching colds.[4]

It is possible for A to be negatively correlated with B even though A, by itself, might be a positive causal factor for B. Here Martin considers the case of a negative correlation between arteriosclerosis and consumption of birth control pills. Although birth control pills are a positive causal factor for arteriosclerosis, pregnancy is a much stronger causal factor, and by massively reducing the chances of pregnancy, consumption of birth control pills reduces the chances of arteriosclerosis.[5]

A particularly notorious example here is that of the long-noted positive correlation between cigarette consumption and lung cancer. Early suggestions, particularly supported by the tobacco industry, included the idea that bodily precursors to cancer (at an early age) cause people to take up smoking, or that some people had a genetic

predisposition to both taking up smoking and developing lung cancer. In this case, of course, neither seems likely, and there are serious questions about why precautionary action was not taken as soon as the correlation was established.

All these examples demonstrate the need for *reliable testing procedures* for definitively establishing the existence (or non-existence) of causal relations, given the existence of some suggestive correlation. And such procedures do exist, though there can be serious difficulties in applying them in some cases.

Controlled experiments

As Martin points out, a typical pattern of scientific investigation from correlation to causation is nicely illustrated by reference to gastritis and bacteria. In 1979, workers in an Australian pathology laboratory discovered what they believed to be a correlation between inflammation of the gut and the presence of a particular bacterium: every tissue with the bacterium was found to be inflamed. A strong correlation was definitively established in 1983.[6]

The question then arose of whether the presence of the bacteria was causing the gastritis (hypothesis one), the gastritis was causing the presence of the bacteria (creating an environment favourable to them – hypothesis two), or some third factor causing both the gastritis and the presence of the bacteria (hypothesis three).

In 1985 one pioneer researcher consumed a mass of the bacteria and subsequently developed gastritis. But as Martin points out, this was a rather small sample (of only one). And even if 1000 people had consumed bacteria and subsequently developed gastritis this would not have definitively ruled out hypothesis two and hypothesis three. They could still have been exposed to the cause of the gastritis before consuming the bacteria. And this other factor could either cause gastritis, which, in turn, causes the presence of bacteria (Factor C) or it could cause both gastritis and bacterial presence (Factor D). The results would be compatible with all three hypotheses.[7]

The next step in this sort of situation is to try to select a study group randomly from the population of interest. This could be rats or other animals or it could be humans. The experimenters then randomly divide the study group (50/50) into an experimental group and a control group. They give all of the experimental group the bacterium and none of the control group. And they aim to treat both groups in

otherwise identical fashion throughout the course of the experiment. In the case of humans they would give the control group a suitable placebo, so subjects do not know they are not receiving the bacterium. This is called *blind testing*.

It is also important that those organising the experiment and assessing the result, in this case determining the percentage of gastritis in the two groups, don't know who got the bacteria and who got the placebo. Otherwise they may treat the groups differently or they may appraise the results differently. The possibility of fraud also arises without such *double blind* procedures.

Now it does not matter if there is a Factor C or D. With reasonably large control and experimental groups, we can expect that C and D in both groups match those in the general population. Then, if all of the EX (experimental) group developed gastritis and only 5% of the K (control) group did, this suggests that (at least) 95% of the gastritis in the EX group was caused by the bacterium.

It does not matter whether we exclude pre-existing gastritis and/or bacteria. If 5% of the general population has both, so will 5% of each of our groups. We should still see a difference between the two groups when the other 95% of the EX group get the bacterium.

Suppose all were initially free of infection and the results were 49% gastritis (EX) and 28% gastritis (K). It then seems that bacterial infection raises the chances of gastritis from 28% to 49%. But we need to calculate the relevant margins of error, assuming a confidence level, and to check for overlap. With a confidence level of 0.95 and a sample size of 100 (EX) and 100 (K), we apply standard error (SE) for a proportion; –1 SE proportion = square root [(f)(1 – f)]n]; for EX = square root (0.5 x 0.5/100) = SR 0.25/100 = SR 1/400 = 1/20 = 0.05. 2 SE [95% confidence level] = 0.10. For K = SR(0.28 x 0.72/100) = 0.045; 2 SE = 0.9

So we're looking at 49% plus or minus 10% = 39–59%; and 28% plus or minus 9% = 19–37%; There is no overlap, and the strength of the causal factor is 59% – 19% = 40%; 39% – 37% = 2%. A positive causal factor with 2% strength is very weak. A positive causal factor with 40% strength is quite strong. A small sample size makes it difficult to assess strength. Nonetheless, the absence of overlap in this case demonstrates a statistically significant result and quite good evidence for the existence of a causal relationship between bacteria and gastritis.

The actual experiment, involving human volunteers, provided good evidence for bacteria as a strong positive causal factor for gastritis,

and led to new cheap and efficient drug treatments for what was previously believed to be a stress-related condition.[8]

This example highlights the problems of using human subjects in unpleasant or potentially dangerous experiments. Fully informed and free consent is obviously crucial.[9] If we are looking at ill people involved in clinical trials of possible treatments, it will probably be crucial to run the new treatment against the best available existing treatment, rather than allowing some to go untreated. (Of course, those receiving the new drug could then be untreated or worse.)

Typically, rats (or other animals) will be used before humans, or instead of humans (where the experiment is expected to lead to serious harm or death). But this too raises serious ethical and technical issues. Are we justified in damaging or killing rats to test the toxicity of possibly quite unnecessary products? How do we ensure that they don't suffer needlessly? Are we justified in extrapolating from humans to rats and back again?

In many cases tests for toxicity or carcinogenicity involve estimation of (possible) lifetime doses for humans exposed to such agents (chemical compounds, electromagnetic fields or whatever): for example, 100 grams of compound X. Rats have one-hundredth of human body weight so we should feed them one gram of the compound. But if we anticipate relatively low levels of the effect (low levels of causal efficacy of the agent), this would involve using large numbers of rats in order to achieve a statistically significant effect, so in practice the dose is increased to 10 grams. We can infer results of normal patterns of exposure from such tests providing the cause–effect relationship is more or less linear. Problems arise where there is a threshold effect: where normal amounts of the agent have no effect but higher ones have a big effect. A possible example here, cited by Martin, is foetal alcohol syndrome. One study suggests that more than three drinks a day are necessary for any of 16 problems associated with the syndrome, but above three drinks such problems proliferate.[10]

A more general possibility is that the cause–effect relation is non-linear, with more cause leading to more effect at all levels. But it could also be that above a particular level a small increase produces an increasing effect. If there is a substantial increase in cases of the effect – in the experimental animal population – with large doses and if we have some reason to expect linearity, then we can conclude that a small dose increases the risk to humans. The crucial question is the extent of the increased risk as against the possible benefits of use of the agent in question.

Prospective and retrospective tests

There are many types of cases where controlled experiments are not possible. Planes cannot be crashed or cars with people in them, large earthquakes cannot be triggered in urban areas. In these sorts of cases scale models can be built, including models of human bodies and such models tested to destruction. Computer models can be built and tested in virtual reality.

In other sorts of cases involving threats to human life and health, including the crucial case of cigarettes and human illness, observational alternatives to experimental methods can be used. These are called prospective and retrospective studies. The basic idea of *a prospective design* is to find two groups of subjects, on average, similar in every feature except for the fact that all members of one group exhibit the suspected causal factor (for example, smoking) and all members of the other do not. Both groups are then followed into the future, and at some point they are compared for the effect variables under investigation (presumably no encouragement is offered to continued smoking, and hopefully strong discouragement).

There are a number of issues here. There are practical problems in finding appropriately homogenous populations or in matching groups for all possibly relevant variables. The major potential cause under investigation may be correlated with other potential causal factors. The 24-year long US National Institute of Health Framington study on the causes of heart disease and other illnesses (looking at 5127 individuals over the period 1950–74), for example, found smoking, a likely causal factor for coronary heart disease, positively correlated with another possible causal factor, coffee drinking. Investigation of the effects of smoking therefore required comparison of non-smoking coffee drinkers with non-smoking non-coffee drinkers. This showed no significant differences between the groups. Investigators could also have looked at smoking coffee drinkers and smoking non-coffee drinkers, but this would have raised problems of possible toxic interaction of smoking and coffee drinking as cause of coronary heart disease (CHD).

Ideally, the study should have aimed to match the incidence of the suspected third factor in the control group with that in the experimental group at the beginning.

> Having obtained randomly selected samples of coffee drinkers and those not drinking coffee, the Framington investigators could have examined

> both samples for smokers. Then, by randomly eliminating some smokers from the sample of coffee drinkers, they could have ensured that there would be the same number of smokers in both groups. Thus, the effect, if any, of smoking would be equalised in the two groups.[11]

But as Giere also observes:

> a drawback of matching samples for suspected third factors is that this strategy does not protect against interactive effects. If coffee drinking contributes to CHD only for smokers … then the EX group would tend to exhibit more cases of CHD than the control group with an equal number of smokers. Investigators could be misled into thinking that coffee by itself is a positive causal factor.[12]

Things get increasingly complicated when we consider greater numbers of potential third-factor correlations, and interactive protective effects.

So the basic principles of a prospective study are:

1 find a homogenous population or collect masses of data in a non-homogenous population relating to possibly relevant variables;
2 distinguish those with and without the supposed cause and match groups;
3 follow both groups long enough for the effect to appear;
4 compare percentage of effects amongst possible cause group with percentage of effects amongst non-cause group.

Such prospective studies can be done either in *real time*, as with the Framington study, or *after the fact*. In the first case we find a group with the cause before they show the effect and follow them possibly for many years. This can be costly and difficult. In the second, we find a group with the cause and then check for the effect. In this case no follow up is required. But if we choose groups with and without the cause only after the effects have shown up, it is less likely those groups will be representative of the population as a whole. Some smokers will have died of heart disease already, so the smokers will look as if they have less heart disease than they really do.

Giere cites this particularly compelling real-time prospective study relevant to the issue of smoking and disease:

> In the mid-1960s the [US] National Cancer Institute sponsored a prospective study that enrolled more than 400 000 men between the ages of 40 and 80 … the investigators matched 37 000 smokers with an equal number of non-smokers [in terms of] age, race, height,

> nativity, residence, occupational exposure to potential carcinogens, religion, education, marital status, alcohol consumption, amount of sleep, exercise, stress, use of drugs, current state of health and history of illness. Every pair of men exhibits the same values of all these different variables, but in each case, one was a smoker and one a non-smoker. After three years and 2000 deaths, the rate of death from all causes was twice as great for the smokers as for the non-smokers. Nearly half of all the deaths were due to heart disease, with the rate for smokers being double that for non-smokers … the death rate from lung cancer for smokers was nine times that for non-smokers.[13]

Another possibility is a *retrospective study*. In this case, we begin with a sample of subjects that already have the effect and look back in time to isolate a suspected causal factor. An example here, cited by Giere, is a British study (by the UK National Case Control Study Group) which looked at all women (younger than 36) diagnosed as having breast cancer between 1 January 1982 and 31 December 1985, in 11 different areas of England, Scotland and Wales: a total of 1049 women. 'For each case in the [final] sample of 755, the researchers selected, at random, one woman as a control from amongst the list of patients of the same physician as the corresponding case. The control subjects [were matched with the experimental subjects], except, of course, they had to be free of any symptoms of breast cancer.'[14] Both groups were then compared for use of oral contraceptives. 'Amongst those with breast cancer, 470 of 755 or 62% had used oral contraceptives for more than four years. Amongst the controls, 390 of 755 or 52% have used oral contraceptives for more than four years.'[15]

We see here the basic principles of retrospective study:

1 collect a sample of those with the effect and match them with a similar sample without it;
2 1ook into the past of the two groups to see whether the group with the suspected effect had more exposure to the suspected cause;
3 compare percentage of causes amongst effect group with percentage of causes amongst non-effect group.

Retrospective studies have the same sorts of problems as after-the-fact prospectives, as well as special problems of their own. As Giere says:

> the data they yield allow no estimate of the effectiveness of a causal factor. Effectiveness is defined in terms of the percentage of the population that would experience the effect depending on whether all or none

> had the cause. The frequencies of the effect in two samples, as in a prospective study, may thus be used directly to estimate the effectiveness of the causal factor in the population [how many people are dying as a result of smoking, for example]. Retrospective studies, however, give you the frequency of the cause in groups with and without the effect. There is no way to use these frequencies to estimate the effectiveness of the causal factor. Knowing that 62% of women younger than 36 diagnosed for breast cancer had used oral contraceptives for more than 4 years tells you nothing about the percentage of women using oral contraceptives for more than four years who will get breast cancer.[16]

Also:

> any observed statistically significant difference between the two sample groups may all too easily be merely a refection of bias in the selection process that produced the samples. There is also the possibility of correlations existing in the population that, even with random sampling, could lead to a statistically significant difference in the sample groups even if here were no causal connection between the variables being studied.[17]

Fallacies of statistical reasoning

Having thus briefly considered the statistical reasoning involved in testing simple statistical claims (relating to proportions), claims relating to correlations and simple causal hypotheses, we can now see various potential areas of confusion and fallacy.

1 First of all, we need to consider the way in which the concepts used in the relevant percentages and proportions have been operationalised. We are regularly bombarded with statistical claims involving terms that are so imprecisely defined that the use of some precise figure in such a claim is meaningless. Just as often we are presented with statistical claims requiring evidence that is practically or logically impossible to obtain.
2 Second, there are problems in ensuring samples are representative of populations. As Hurley points out, it is no good a quality controller for a manufacturing firm checking every tenth component on a conveyor belt if the components are not randomly arranged on the belt. 'As a result of some malfunction in the manufacturing process it's quite possible that every tenth component turned out perfect and the rest imperfect.'[18]
3 The randomness requirement presents particular problems where the population consists of people, with phone polls, for example, effected by different patterns of work at different times of day. Not

> everyone has a phone or an address appearing in a city directory. And, as noted earlier, people's answers will not necessarily be error-free or truthful.

There are opportunities for error or fraud in proper application of the procedures for collecting and interpreting the data (considered above). Particularly insidious is the case of corporations carrying out numerous trials of the causal efficacy – or safety – of new products, but publishing only results favourable to their own economic interests. As should be clear, even if a product is useless or dangerous, enough trials will, by chance, eventually yield results that make it appear useful or safe.

As Chalmers notes, the new generation of antidepressants were originally billed as

> effective, safe and non-addictive … Now [after billions of dollars worth of sales worldwide] … some patients and doctors claim they are of questionable efficacy and can induce suicidal thoughts … It would be easier to judge which side was right if all the relevant information about the drugs were publicly available. But … the law does not oblige companies to disclose the findings of their research on licensed medicines and scientists, doctors, patients and … public organisations have no legal right to inspect the evidence that led regulators to licence drugs [let alone other results not made available to regulators].[19]

Another consideration here is the use of *surrogate endpoints* in clinical trials. A new drug is tested for a comparatively short time and is found to produce some apparently desirable result, for example, reduction of high blood pressure. On this basis, it is approved by relevant government bodies and widely prescribed. As Moynihan points out:

> the trouble is, high blood pressure itself is not a disease that needs treating; it is … only a risk factor for stroke or heart disease. Preventing these diseases is the real aim of drug therapy. In order to measure the really important effects of the drug it is necessary to conduct long-term studies which will measure the number of strokes or heart attacks which the drugs prevent in patients, rather than whether they simply lower their blood pressure.[20]

He cites the case of the drug flecainide, found in brief clinical trials to reduce irregular heartbeats in people at risk of cardiac arrest. It was assumed that if the drug reduced irregular heartbeats then it would prevent heart attacks, and it was widely prescribed. A belated large-scale trial showed that in fact the drug was causing cardiac arrest in those with less serious symptoms of irregular heartbeat. Health policy

researcher Moore estimated that 50 000 patients had died from taking flecainide and the related encainide before they were withdrawn. As Chalmers says, 'by 1990, more than a decade after these drugs were introduced, it has been estimated that they were killing more American every year than died in action in the Vietnam war'.[21]

Gigerenzer highlights a further problem area concerning the way in which the results of clinical trials are presented. He cites the case of a press release from the West of Scotland Coronary Prevention Study of the effect that 'People with high cholesterol can rapidly reduce … their risk of death by 22% by taking a widely prescribed drug called prevastatin sodium. This is the conclusion of a landmark study presented today at the annual meeting of the American Heart Association.'[22]

This information presents a potentially misleading picture of the efficacy of the new drug to the statistically naive. As Gigerenzer says, 'studies indicate that a majority of people think that [this means that] out of 1000 people with high cholesterol 220 of these people can be prevented from becoming heart attack victims [through the use of this drug]'.[23] The clinical trial actually showed that out of 1000 people taking prevastatin over five years, 32 died compared to 41 in a placebo control group.

The relative risk reduction was indeed 22% and the drug manufacturers were happy to use this figure in promoting their product. However, the absolute risk reduction, probably of more interest to both potential users and governments subsidising public medicine, is 0.9%. 'Prevastatin reduces the number of people who die from 41 to 32 in 1000. That is, the absolute risk reduction is 9 in 1000.'[24]

Furthermore, the number of people who must participate in the treatment to save one life is 111, because 9 in 1000 deaths – which is 1 in 111 – are prevented by the drug. This could be a substantial cost to a public health system, with the money better spent elsewhere. Most obviously, there is the possibility of much cheaper, less dangerous non-drug-based means of reducing high cholesterol.

Recognising that good clinical trials are expensive and difficult undertakings, we must recognise also that the wealthy private companies that can afford to fund such trials are highly unlikely to fund trials of already available – non-patentable – alternative (including non-drug-based) treatments from which they can derive no profit. Nor are they likely to develop drugs to treat the illnesses of the poor (however widespread and serious) because the poor cannot afford to pay inflated prices for patented drugs.

Legal issues and the precautionary principle

There are many legal issues arising out of consideration of Type 1 theory testing. Here we focus upon two particularly important and closely related issues concerning correlation and causation. The first concerns the centrality of the requirement of proof of causation as the basis for culpability in criminal law and liability in the law of negligence. Even 'regulatory' laws, setting permissible standards in relation to workplace safety, food safety and pollution, and aiming to thereby avoid possible damage to people or the environment in the future, have generally only been enacted as a result of proof of the damage-generating potential of particular practices in the past.

We have seen some of problems in definitively establishing causation; it can be extremely difficult, costly and time consuming, taking decades rather than months. By the time such definitive proof of causation is available, terrible damage can have been inflicted upon people and the environment, on a vast scale. It is of little benefit to the victims that their deaths contribute to eventual legal action years down the track. At the same time, we have also seen that evidence of correlation, along with other circumstantial evidence, can point very strongly towards particular patterns of causation long before any such 'definitive' evidence becomes available.

There are many well-known examples. Awareness of the dangers of asbestos, for example, goes back to ancient times, at least in areas of large accessible asbestos deposits. In the West, in 1898, a woman factory inspector reported 'injury to bronchial tubes and lungs' of workers 'medically attributable' to asbestos in the workplace.[25] This was followed by a succession of similar observations of lung disease amongst asbestos workers by doctors and factory inspectors. But it was only in 1931, following a British government inquiry, that the first asbestos dust control regulations were introduced, along with medical surveillance of workers and compensation arrangements.

In the 1930s and 40s reports of lung cancers being associated with asbestos appeared in the US, UK and German medical literature. A British report of 1953 found a rate of lung cancer 10 times greater in asbestos workers than in the general population. And a South African report of 1960 found all but 2 of 47 cases of the rare cancer mesothelioma involved prior asbestos exposure, in some cases involving only very short periods of exposure.

Yet as Gee and Greenberg point out, the UK asbestos regulations

of 1931 'were only partially enforced', with only two prosecutions between 1931 and 1968. Updated regulations in 1969 failed to consider the cancer hazards. And only in 1998 following a World Health Organisation report of 1986, finding all types of asbestos to be carcinogenic, with no known safe level of exposure, did the British government adopt a ban on all forms asbestos. A ban in the European Union allowed continued use of asbestos till 2005. 'Meanwhile, the annual UK cancer death rate from mesothelioma, a lung cancer from asbestos, is estimated by the Health and Safety Commission to be around 3000 deaths per year and rising.'[26]

It was only in the early 1980s that asbestos was banned for most uses in Australia, once the authorities had 'strong epidemiological evidence' of the dangers. As Deville and Harding note:

> a consequence of the lack of precautionary action is that many people have died from asbestos related disease, there have been large claims against corporations involved with asbestos mining and manufacture and Australia has the highest rate of mesothelioma in the world.[27]

We can see how early reports, although perhaps less than 'definitive' – insofar as they involved small samples, retrospective studies and after-the-fact prospective studies – were nonetheless highly suggestive of causation. The first of the modern reports (that of factory inspector Lucy Deane in 1898) even pointed to the precise nature of the mechanism involved, identifying the 'sharp, glass-like jagged nature of the particles'[28] revealed by microscopic examination.

Most important was the nature of the damage involved here: a serious threat to life and health for vast numbers of people. Given this sort of combination, of possible or likely causation with intensity and scale of possible or likely damage, it seems clear that radical legal action should have been taken decades earlier.

There are many similar stories of failure to take action in time to save health, life and the natural environment. It has become increasingly clear that waiting for definite proof of causation can be the height of social irresponsibility in many cases and that it is vitally necessary to apply some sort of strong *precautionary principle* to avoid such disasters in the future.

The basic idea here is clearly explained by Deville and Harding:

> Where there are threats of serious or irreversible damage [to people or to the environment] lack of full scientific certainty should not be used as a reason for postponing moves to prevent [such damage][29] … Under

> the precautionary principle it is the [potential victims] rather than those whose actions may impact upon [them] that [are] given the benefit of the doubt.[30]

And only with strong evidence of a big balance of real social benefits over costs are new projects allowed to proceed.

It is crucial here to put the onus of proof of safety (or social benefit significantly outweighing safety risks) onto any organisations or individuals preparing to institute such new projects, market some new product or introduce some new technology. Potentially safer alternatives must always be seriously considered before any such project is allowed to proceed. And those in a position to approve such new developments should try to apply the most precautious project management in the early stages. Ongoing and effective monitoring of all potentially dangerous consequences of projects of all kinds is an integral part of the program, with strong action taken at the first signs of trouble.

> Many case studies indicate the value of thorough, long-term monitoring. While for asbestos, benzene and PCBs evidence was accumulating of the adverse health effects ... no role was then played by systematic monitoring.[31]

Deville and Harding argue for legislation leading to the establishment of environmental regulators to require all such bodies to apply the precautionary principle in decision-making. They call for the banning of any – potentially – seriously dangerous activities before scientific proof of damage, for 'reverse' lists of safe substances to replace lists of known hazardous substances banned from particular uses. All substances should be assumed to be harmful until there is good evidence that they are not. 'The long-term effect of this listing is that the onus is then on the developers, users and disposers of particular substances to prove that they should be on the list. Reverse lists should be linked to [extended principles of strict and absolute liability].'[32]

Harremoes et al emphasise the importance of extending regulatory appraisal beyond the most 'straightforward and direct impacts', to 'as wide a range of conditions and effects as can reasonably be anticipated'.[33] They also argue that we need to take 'account of the potential irreversibility of actions, even if the consequences might not be known'. [34] And they cite the examples of halocarbons (dissolving the ozone layer), PCBs (accumulating in the food chain) and methyl tert-butyl ether (as a potential serious-illness causing petrol additive) as

artificial chemicals whose 'very novelty' should have been taken as a warning sign long ago.

> Enough was known at the outset regarding their persistence in the environment to serve as another warning. They would also readily disperse to become ubiquitous throughout the physical environment … it could have been reduced from the outset that if these substances were released into the environment and if a problem subsequently developed, it would take many years for both them and the problem to 'go away'.[35]

They emphasise also the need for regulators and others to search out and address 'blind spots' and gaps in scientific knowledge relevant to issues of potential serious and irreversible harm. 'For halocarbons the chemical mechanism for depletion of stratospheric ozone was identified in the prestigious journal *Nature* in 1974. Nevertheless, this did not prevent regulatory neglect until first empirical evidence of causal effects became available.'[36] They focus upon the need for interdisciplinary risk assessment (including use of lay and local knowledge as well as specialist expertise), for taking account of 'real world conditions', for proper systematic scrutiny of claimed pros and cons, and, above all, for regulators effective independence from economic and political special interests.

Developers of all kinds will (and do) argue that further moves towards precautionary policies will radically slow the pace of technological innovation. And it is true that we must always consider the costs of failing to develop new technologies, as well as the potential costs of the technologies themselves. In particular, some risks are justified in the interest of saving human life and restoring the health of the environment. But genuine technological progress, as distinct from profit-driven innovation, should be predicated upon ensuring safety. The disasters cited in this chapter are enough to demonstrate the need for truly responsible scientific caution.

Social science

Before leaving the issue of the precautionary principle, it is important to briefly note its significance in relation to other areas of law reform. In particular, the work of Wilkinson and other researchers has focused attention upon an increasing body of empirical evidence relating to the social causes of crime, accidents and ill health. Amongst other evidence, Wilkinson refers to a strong correlation (above 0.72, P = 0.01) between homicide rates (0–20 per 100 000 population) and the share

of total household income received by the least well off 50% of the population (17–23%) across 46 US states for which data is available. And he refers to data from more than 30 other countries establishing a strong correlation between inequality and violent crime.

It turns out that greater income disparities, around the world, are correlated with reduced life expectancy, particularly for those at the lower end of the income scale, and greater incidence of death from illnesses, alcohol-related conditions, traffic accidents and injuries. And an older study of 192 metropolitan areas in the United States between 1967 and 1973 found 'clear relationships between most of the major categories of crime and the size of the "income gap" between the incomes of the poorest 20% of the population and average incomes in each area'.[37]

These (largely correlational) studies do not provide definitive proof that inequality causes violent crime (and serious illnesses and accidents). It could be suggested that the causal arrow goes the other way. Or that some third factor is causing inequality, accidents and crimes. Undoubtedly, there are multiple complex interactions of different variables in this area. However, there is a mass of other relevant evidence, nicely explored by Wilkinson, to support the hypotheses of inequality as a major causal factor in all of these areas. In particular, evidence from many different sources suggests that stress in the family and at work, associated with feelings of general powerlessness (destruction of immune systems, frustration and anger), is the key *result* of income inequality, which, in turn, becomes the *cause* of accident, illness and crime.

As Wilkinson notes, the fact that

> links between crime and income inequality to some extent parallel those between health and inequality, is highly indicative of the channels through which health is affected. It not only provides independent confirmation that income distribution has important psychosocial effects on society, but shows that the effects are consistent with the view that wider income differences are socially divisive. Indeed, there are suggestions that they undermine the legitimacy of the society's institutions more widely.[38]

Coupled with substantial empirical evidence of the radical failure of criminal law to effectively reduce crime (through traditional policing and punishments – see chapters 18 and 19) and of tort law to effectively reduce 'accidents' or properly compensate victims, this data provides a very strong 'precautionary' foundation for a radically different approach.

Causation in populations

While major social-structural reforms could radically reduce accidents and injuries (through reducing social inequality and applying a strong precautionary principle to release of potentially dangerous products), it is important also for the law to catch up with the sort of understanding of causation achieved through the experimental studies considered so far.

A basic principle of the tort of negligence is the requirement of the plaintiff to establish, among other things, that the negligence of the defendant caused them injury or other detriment. Tort textbooks still suggest that 'common sense' is all that is required to establish whether or not a particular act or omission has caused damage to the plaintiff in any particular case. However, we have already seen good reasons for doubting any such idea.

Causation might indeed have been a relatively simple matter in earlier periods of development of the common law when it was merely a matter of establishing whether or not A really did hit B hard enough to produce the relevant injuries, or whether or not A's pig escaped and ate B's cabbages. But as we have seen, establishing the causation of lung cancer and heart disease in years gone by has been a rather more complex issue. And establishing the health effects of the use of mobile phones or of mass consumption of genetically modified foods or new drugs controlling blood pressure will similarly depend upon large-scale epidemiological investigations and complex controlled experiments.

Where relevant research has already been carried out, and relevant data already exists, lawyers need a basic knowledge of scientific research methods (such as that provided in this chapter) to begin to make sense of the material. But there are serious problems of the courts previous inability to properly understand and apply such knowledge.

The fundamental problem concerns the difference between causation in individuals and in populations. The law's basic model of causation is one involving direct perception of the causal activity (the exercise of the causal powers of) individuals. But the sorts of tests we have considered yield results in terms of increased percentages of particular effects in particular populations as a result of exposure to particular causal factors.

In many cases the harm inflicted by the pursuit of corporate profit registers in just such an increasing percentage of cases of some illness

in a large population: more cases of heart disease, or lung cancer or asthma, or whatever. It is quite possible, in some such cases, to establish the percentage involved, without being able to trace the actual pathway of (mechanical) causation in particular individuals. That is to say, we might have solid evidence (from observational studies) to prove that, for example, 60% of the lung cancer cases in this group of workers have been caused by their exposure to asbestos in a particular workplace. But we cannot 'prove' the causation in any particular case. The other 40% of cases in this population will have different causes and we can't show which is which.

The standard of proof in tort cases makes direct reference to the 'balance of probabilities', which means a more than 50% chance of causation. In this case, this requirement would seem to be satisfied; in any individual case the chance is more than 50% asbestos causation. But it is quite possible that, while causing a significant percentage of cases of a particular illness, a defendant's negligence actually causes less than 50% of all such cases, meaning that there will be a less than 50% chance of causation in any particular case. But then it seems that in law the defendant in question gets off completely, even though they might have knowingly (or negligently) killed thousands of people.

The idea that a defendant corporation (or its controllers) should escape liability for increasing the prevalence of a particular cancer in a population simply because they remain responsible for less than 50% of cases overall is absurd.

If 15% (3000 people) of population A, exposed to chemical X, succumb to cancer Y while only 10% of population B, not exposed to X, but in all other ways identical, succumb, then it is true there is a less than 50% chance that the cancer of any particular sufferer in A was caused by chemical X. Reference to appropriate observations of the two matched populations (even with properly statistically significant results) fails to prove causation 'on the balance of probabilities'. But if executives of chemical company C were responsible for such exposure, despite their having prior access to information from earlier studies showing the cancer-causing powers of X, then they are knowingly or negligently responsible for killing or seriously damaging 1000 people.

It is easy enough to see 'logical' ways in which the law could deal with such a situation (which never should have been allowed to develop in the first place). While there is a greater than 50% chance that any individual victim considered in isolation contracted their cancer from something other than chemical X, if we consider two victims together,

there is only 2/3 x 2/3 = 4/9 chance that both got it from something else, or a greater than 50% (5/9) chance that one got it from chemical X.

We could say that for any pair of two victims there is a greater than 50% chance that one of them got their cancer from chemical X; therefore, the company should pay each victim half of what they would have paid with direct 'individual' proof of causation. Or we could say that the company caused a third of all cases in population A so they should pay each victim a third of what they would have paid with such direct proof.

The development of 'market share liability' in the case of *Sindall v Abbott Laboratories* 26 Cal 3d 588; 607 P2d 924; 163 Cal Rptr 132 (1980) in the United States represents significant progress in this area. This case involved the widespread damage caused by the synthetic hormone diethylstilbestrol (DES), marketed worldwide to pregnant women between 1940 and 1971 to prevent miscarriage. There was no doubt about the negligence of the manufacturers who had failed to carry out basic tests of the material and ignored the damning results of other tests (which showed that it actually caused miscarriages as well as cancers). The problem for the plaintiffs lay not in proving their injuries were caused by DES but in establishing which of the around 200 manufacturers had made the DES they had consumed.

The Supreme Court of California held that each manufacturer 'be held liable for the proportion of the judgment represented by its share of the market unless it [demonstrated] that it could not have made the product that caused [the] plaintiffs injuries'.[39] They therefore dispensed with the requirement for a necessary mechanical linkage in every individual case. It is only a small logical step from such market share liability to a general principle of liability in proportion to increased percentage effect produced in relevant populations. Hopefully, developing awareness amongst responsible lawyers of the true nature of causal relations, and the scientific means for establishing their existence, will accelerate the process.

Discussion topics

1. What are causal hypotheses? How are they tested? Include reference to prospective and retrospective experimental designs.
2. What are fallacies of statistical reasoning? Include examples (not just from this chapter).
3. Explain the precautionary principle and causation in populations.

Debate topics

1. Ideas of causation in common law should be guided by science, not 'common sense'.
2. Serious application of the precautionary principle would put an end to scientific and economic progress.

Additional resources

S. Beder, 'Scientific Controversy: Dioxin', in *Global Spin*, Scribe Publications, Carlton North, 1997, pp 141–60.

A. Chetley, *Problem Drugs*, Stirling Books, Australia, 1995, ch 19 on DES.

R. Graycar and J. Morgan, *The Hidden Gender of Law*, Federation Press, Sydney, 2003, pp 334–6 (section on dangerous products and dangerous drugs).

E. Handsley, 'Market Share Liability and the Nature of Causation in Tort' (1994) *Torts Law Journal* 24 at 24–44.

S. Mann, *Economics, Business Ethics and Law*, Lawbook Co, Sydney, 2003, chs 17 and 18.

D. Weatherburn, 'What Causes Crime? (section on poverty and unemployment), <http://www.lawlink.nsw.gov.au/bocsar1.nsf/pages/cjb54text>.

R. Wilkinson, *Unhealthy Societies*, Routledge, London, 1996.

Chapter 7

THEORETICAL HYPOTHESES

What are theoretical hypotheses?

In chapter 5, references were made to Type 2 theories, but no details of the nature of such theories were provided. This chapter provides a concise introduction to such Type 2 theories.

To understand Type 2 theories, we must first recognise that the further progress of our understanding of the nature, properties, powers and behaviours of things requires us to go beyond directly observable and measurable phenomena. In many cases, there is, anyway, nothing in the world as directly perceived that offers any kind of explanation for observed properties of things. If we want to find out why and how particular observable things have the properties they do have (why metals expand when heated, why rich people live longer), why and how they are able to exercise the particular observable causal powers they do exercise (how salt dissolves in water, how asbestos causes cancer), or why and how particular events – perhaps events in the remote past – came about (how the earth came into being, how mountains were formed), we typically have to consider internal structures and external relations of such things that are either too small, too quick, too big, too slow, too long ago, too far away, too complex or too deeply buried for any kind of direct observation. In this sort of case, we are completely dependent – in the first instance – upon *imagination* to construct hypothetical models of such unperceivable structures and mechanisms. Later, under the guidance of such models, we can sometimes construct instruments, extensions to our sensory systems, capable of rendering such structures and mechanisms more or less directly perceptible.

We can see how *analogy* plays a central role in the construction of such higher-order Type 2 theories. More specifically, we can see the importance of the sort of 'creative' analogising considered in chapter 1, which takes account of differences as well as similarities, in extending human understanding beyond the limits of what is directly perceptible and controllable.

The key idea here is that we find something similar to the phenomenon in need of explanation, but whose causation has already been established, typically through direct observation. We then reason back to similar, hypothetical causes for the phenomenon of unknown causation, allowing for the likely difference of scale involved.

P1 We know that object d (the mountain) resembles object x (the molehill) in respects e, f and g (it is made of similar earthy materials, has a similar location, rising up from the surface of the earth and a similar – conical – shape). It also differs from object x in respects h and j (it is much bigger and older).

P2 We know that object x (or objects of type x) is a product (or are products) of the action of causal agent y (or agents of type y) acting upon some pre-given material z in circumstances a, b, c (this molehill has been brought into being by a mole piling up earth and stones from its underground tunnelling last night).

C We therefore formulate the hypothesis that the object d is (possibly) the product of the action of a causal agent j, similar to y, but differing from it in ways related to h and j (the mountain has been brought into being by a giant mole tunnelling under the earth long ago in the dreamtime).

Here we have constructed a sort of *model*, a model of an unknown mechanism (the giant mole). Our model is also a theory, a theory to explain an otherwise inexplicable phenomenon by reference to a particular sort of mechanism (giant moles created the mountains). In this case, the reader's first response is to say that this is not science but rather a pre-scientific animistic speculation. And such thinking is indeed characteristic of pre-scientific cultures. But it is important to see that a precisely similar form of reasoning has played a central role in the development of modern science, providing the first insights into the (unobservable) causal mechanisms operative in many different areas.

Consider the following cases:

- Democritus and Leucippus' atomic theory and its modern development as the kinetic theory of matter;
- Harvey's theory of the circulation of the blood;

- Darwin's theory of natural selection and Mendel's theory of genetics;
- Wegener's theory of continental drift and the modern theory of plate tectonics;
- Van Helmont and Pasteur's bacterial theory of disease;
- Einstein's general theory of relativity (his gravity theory).

In every case we are concerned with theories at the heart of specific disciplines of modern science (including the modern explanation of mountain formation). And in all of these cases it is easy to see the central role of analogical reasoning in the development of the theories in question, extending human thinking into areas too small, too big, too slow, too complex or too distant in time and space for direct perception.

What distinguishes the science from the pre-science has to do with the nature of the assessment procedures applied to such theories, rather than the mechanism of creation or the basic structure of the ideas.

- In the first instance, such hypotheses must be *internally consistent*, *possible* and *plausible*. In the case of the mole theory, we should ask such things as: Does what we know of a small mole's physiology present any problems for our hypothetical large mole? Can the design be effectively scaled-up? Would such a scaled up version be strong enough to shift the sorts of weights of material involved? Could any living creature do so? How would these moles have lived? What would they have eaten?
- Second, we need to consider the consistency of the new theory with other theories we already accept as true. Does the theory generally fit in with what we know of biology and geology? Could the moles have lived in the sorts of environments we believe existed in the remote past? A theory that does not conflict with our background beliefs is said to have the virtue of *conservatism*.
- But most important as far as science is concerned is that the theory be *testable* and *capable of generating significant new knowledge* of matters of empirical fact, over and above providing a convincing *explanation* of the (problematic) facts that originally motivated its formulation. These two considerations, of testability and of new knowledge generation, are closely related.

As far as testability is concerned the issue is straightforward. As Schick and Vaughn observe:

> Since science is the search for knowledge, it's interested only in those hypotheses that can be tested – if a hypothesis can't be tested, there is no way to determine whether it's true or false.[1]

And a theory's capacity to generate new knowledge refers to the ability of the theory to generate new and unexpected predictions of observable facts (deduced from the theory, along with specific initial conditions – that is, states of the world), which turn out to be *confirmed by observation*. Thus, although the theory itself refers to entities or processes that are not, at the time, directly observable, it can still generate predictions relating to things that are so observable.

Atoms were not directly observable at the start of the 20th century. But the atomic theory turned out to be testable by reference to the (observable) Brownian motion of pollen grains immersed in water. Einstein developed a model of molecular motion, based upon consideration of the behaviour of solute molecules in a solvent, which yielded quite precise predictions as to the behaviour of such particles.

To be suitably unexpected and 'independently' testable such predictions must relate to types of things different from those the theory was originally formulated to try to explain. It would be no good to derive a prediction from the mole theory to the effect the more mole-created mountains will be found on hitherto unexplored continents. We would expect to find more mountains whether or not they were created by moles.

Such predictions must be different from predictions that can be generated by other theories, either already accepted as true or competing with the new theory for our allegiance. If we already have a theory that predicts the existence of giant moles in the past, then a prediction derived from our theory to the effect that we might hope to find the fossilised remains of such moles does not provide a suitable test of the theory. Were we to discover such remains, such evidence would fail to favour one theory against the other (though it could be said to offer some support to both).

Another, possibly more suitable, prediction might be the existence of huge tunnels beneath mountain ranges. If mountains were indeed the products of the tunnelling activities of giant moles, then we might reasonably expect to find such further traces of their activities, perhaps including fossil mole bones or coprolites inside the tunnels.

And, of course, if we did find such tunnels, this would be a substantial addition to our knowledge of the world; the theory would have demonstrated its value as a knowledge-generating tool and in the process it would have received substantial empirical confirmation.

If the theory were sufficiently rich to suggest other explanations and tests (in other areas) this would be a further point in its favour. Perhaps the moles are implicated in the creation of other geological phenomena, or other things altogether.

This example also demonstrates some of the problems involved in testing Type 2 scientific theories. For it shows how the progress of science is dependent upon the development of increasingly sophisticated *technology* – in this case, of drilling equipment or seismic or radar systems capable of detecting and mapping underground passages. Real scientific theory testing has been crucially dependent upon telescopes, microscopes, particle accelerators and a host of other sorts of apparatus.

The example also shows that while it is possible to provide strong verification for a new theory, it is difficult to provide any such strong refutation. Failure to confirm the prediction, even after long and detailed investigations with powerful technology, does not necessarily show the theory to be false. It is quite possible that the fault lies not in the theory itself but rather in the background assumptions we have had to make in order to render it testable. The tunnels might not have been strong enough to persist down through the millennia, but could rather have been crushed by the weight of material above, or they could have filled up with silt and rocks from underground streams.

Usually, some of our assumptions will indeed be wrong, even with the best theories, leading to failures to confirm some predictions. But as long as some new predictions are confirmed, this tells us we are on the right track. It means that our theory is successfully generating new knowledge in the form of new (observable) facts about the world, and we can continue to work on our background assumptions to try to find out where the other predictions went wrong.

Broadly speaking, as long as verifications keep ahead of refutations, we have what philosopher of science Imre Lakatos calls a 'progressive research programme' and it is rational to push ahead with the theory in question. But once refutations get the upper hand, once the theory ceases to lead the way to substantial new knowledge and rather seems to follow after empirical studies, with continuous modifications

to the theory 'after the event', then we have a 'degenerating' research tradition.

If our theory fails to deliver any such substantial new knowledge despite our best efforts, then, at some point, it's time to think again about the theory itself and start looking for alternative possible causal mechanisms to explain the original inexplicable facts, along with the accumulating anomalous observations.

Of course, we can't just abandon theories that have demonstrated some real explanatory content (through a good record of confirmation) in the past without having something better to replace them. Nor should we necessarily reject theories (like some interpretations of quantum theory) which are currently untestable. As suggested above, the crucial technology for proper testing might not yet be available. Some theories display virtues of 'simplicity', ' scope' and 'fruitfulness' in predicting hitherto unknown phenomena well before the possibility of effective testing of novel predictions.

Scientists originally accepted Copernicus' heliocentric cosmology, despite empirical refutation (and no confirmation) because it was simpler than the preceding Ptolemaic theory. And, ultimately, their support was justified as new technology and advances in other areas of sciences came to provide solid empirical confirmation.

As noted earlier, the theory itself, along with its novel predictions, often functions as a guide for the construction of new scientific instruments. And, ultimately, via such instruments, the originally unobservable underlying entities and mechanisms identified by the theory can actually become (more or less) directly observable, as a particularly powerful *verification of the theory*. But even if they do not become so observable, the theory can still be *strongly supported* via a developing tradition of confirmation of far-reaching novel predictions manifesting the causal consequences of the operation of such underlying entities and mechanisms.

Deductive reasoning in scientific theory testing

So far we have concentrated largely upon non-deductive patterns of reasoning: specifically, the inductive-analogical reasoning involved in Type 2 theory creation or construction. But a deeper understanding of the nature of theory testing involves reference also to some basic patterns of deductive reasoning.

Three basic forms of valid deductive argument are particularly rel-

evant here. Two we are already familiar with, *modus ponens* and *modus tollens*.[2] The third, DeMorgan's Law, is equally straightforward.

P Not (A and B).

C So not A or not B (and vice versa).

These two statements are logically equivalent, so each necessarily implies the other. If it's not the case that A and B are both true (together) then at least one of the propositions, A or B (or both), must be false. And if at least one is false, they cannot both be true.

We have seen how analogical reasoning allows us to construct 'models' of underlying causal mechanisms or structures, responsible for producing particular observed appearances (events, processes, situations). Such models include reference to specific causal agents (electrons, quasars, tectonic plates), to specific causal powers or abilities of such agents (to repel negative charges, radiate galaxies of energy from star-sized volumes of space, change the patterns of the earth's crust), and to the particular circumstances in which such powers are actually exercised or realised to produce some (observable) effect or consequence (the flow of a current, creation of an image on a photographic plate, the formation of a mountain range).

Such causal mechanisms are 'triggered' or become operative in producing observable consequences in particular sorts of situations only. In other situations, their powers remain dormant, are blocked by countervailing forces or produce effects that cannot be distinguished by human observers, leading to no such observable consequences or effects. (Current only flows with a potential difference in a circuit).

In some cases, we will be concerned with mechanisms or processes that might no longer be operative, or no longer be able to be triggered into operation (or operating so slowly in the present as to be invisible). This is often the case where we seek to explain specific occurrences in the past – unique individual events such as the extinction of the dinosaurs or the origin of the universe – rather than specific types of occurrence continuing in the present like the periodic orbits of the planets, the heating of the earth's core, the appearance of antibiotic-resistant bacteria, or earthquakes in California.

In the latter sort of cases, we will be concerned with mechanisms that are presumed to be operative still, such as radioactive decay, biological evolution and plate tectonics. But in both sorts of cases, we

need to be able to deduce predictions of possible observable consequences other than those that originally motivated the production of the theory. And to do this, we must take account of the specific conditions surrounding the mechanism in question: the interaction of the mechanism in question with its ambient environment of other such mechanisms and processes.

In the simplest sort of case, of a mechanism still thought to be operative in the world, and capable of being triggered (or 'liberated') by direct human intervention, we can create the conditions ourselves. For example, in 1919 physicist Ernest Rutherford used a naturally radioactive material as a source of a beam of alpha waves in a vacuum. He focused the beam into a sample of nitrogen gas, thereby transforming some of the nitrogen into oxygen, and demonstrating the power of such particles to transmute one physical element into another. Later investigators used electric and magnetic fields to artificially accelerate such beams of charged particles, to create many new isotopes.

In other cases we will have no such direct control of proceedings. But we can still, hopefully, generate testable predictions by reference to the generative mechanism and its surrounding conditions.

To take a simple and well-known example, suppose that we have formulated the theory that the impact of a substantial asteroid or comet, of 10 kilometres or so across, was responsible for the extinction of the dinosaurs (and half the other species then alive through massive climatic disruption) at the end of the Cretaceous period, about 66 million years ago.

This theory was originally motivated by discovery of a thin layer of the heavy metal iridium – apparently of extraterrestrial origin – in rock deposits of this period. But in order to further test the theory, we must look for some further observable evidence, other than that which suggested the theory in the first place. And we do not have far to look. Once we begin to consider other aspects of the environment of the impact, and other likely consequences of such an impact within such an environment, other possibilities immediately suggest themselves. Most obviously there is the hope that we might find an impact crater of the appropriate size and date, still observable today. (Other consequences might include evidence of particular sorts of climatic disruption at the relevant time.)

In this sort of case, where we cannot actually control the system concerned to exclude interfering or counteracting forces, we have to

assume that such forces are absent. In this case, we assume (or hope) that the impact was on land, rather than in the sea, and that the land in question is still available for observation, rather than having been completely eroded away, covered by deep ocean sediments or 'subducted' beneath a continental plate.

In the following, M is our model or theory, C our initial conditions and P is our prediction.

> M = Comet of type x strikes land at time y (and the land it struck is still visible).
>
> C = A comprehensive search of all land areas is carried out (of a type suitable for detecting comet craters).
>
> P = A crater – of appropriate type – is found.

We now have a definite prediction: M and C together imply P. And if our reasoning is sound, we can then take a second deductive step. We now know the truth of the conditional statement

> If (M and C) then P.

The truth of 'M and C' is sufficient for the truth of P, since 'M and C' implies P.

As we have seen, we want to produce predictions that are quite novel and unexpected. In particular, we do not want to duplicate predictions than can already be derived from other known or accepted theories. For then, confirmation of such predictions would provide support for both theories, in no way favouring our (new) theory. Nor would it provide any genuinely new knowledge about the world.

As Ronald Giere suggests, we can capture this requirement in the idea that

> If (not M and C) then very probably not P.

If the theory (M) is false (but we are otherwise correct about the conditions at the time of our test (C)) then very probably the prediction will not eventuate; we do not expect it to be confirmed under the circumstances in question (C).

One obvious way to satisfy this criterion is to *make the prediction as precise as possible*. The more precise, the more unlikely. Notice that the crater prediction from the cometary impact theory seems to satisfy this criterion quite well, providing that size and date are precisely specified and checked.

Producing such a novel prediction, and establishing that it really is novel, will often be a far from straightforward process. Interesting theories or models will often be complex, and their relations to perceived (or measurable) reality by no means direct. A great deal of cognitive work can therefore be required to produce such a testable prediction.

Consider the search for dark matter or for black holes in the universe. By definition, we are concerned with types of things that cannot be (or can only with great difficulty be) observed directly; similarly with Freudian psychology and its postulation of unconscious mental processes and agencies. Even a theory of such antiquity and power as atomic theory was still rejected by some authorities at the beginning of the 20th century on grounds that no definitive empirical evidence was available.

But this has not prevented committed researchers from devising ingenious tests, and unexpected predictions that really do follow from the theories in question – along with consideration of appropriate (initial) test conditions(as in the examples above).

Once we have such a prediction, we can carry out our experiment or test of the theory. As noted above, in some cases this will actually involve creating condition C capable of triggering mechanism M into producing observable effects. And this will typically involve isolating the mechanism in question as much as possible from the influence of others that might block or obscure its operation.

In other cases, as with mechanisms no longer operative, or large-scale geological, astronomical or cosmological processes, on a scale beyond the reach of human intervention, we have no such direct control. But the principle remains the same. We must take account of the relevant causal mechanism and the (presumed) specific conditions of its operation to guide our search for relevant evidence.

On the one hand our prediction can be refuted: wide-ranging searches fail to reveal any crater of the right size and age. The logic of *modus tollens* (MT) and DeMorgan (DeM) then tells us that we are wrong about our proposed model or about the prevailing conditions, or we have simply failed to carry out our experiment properly.

P1 If (M and C) then P.

P2 Not P.

C1 So not (M and C). (MT)

C2 So not M or not C. (DeM)

Certainly this does not necessarily mean we are on completely the wrong track. We could be right about the basic mechanism involved; we just need to refine our understanding of the conditions. Perhaps conditions are not actually such as to allow the mechanism to operate as predicted; some other forces were operating to counteract it or block it after all. Perhaps the comet landed in the sea and could not therefore produce a crater. Perhaps the crater has indeed disappeared beneath ocean sediment or below continental rocks. Perhaps we have simply failed to search long and hard enough.

On the other hand, our predictions could be confirmed. There is the crater in all its glory, at the tip of the Yucatan peninsula; there is the measurement, precisely as predicted. In this case, *modus tollens* and DeMorgan give us good grounds for taking the theory seriously and continuing to develop it further.

P1 If not M and C then very probably not P.

P2 But P.

C1 That is, not not P (double negation).

C2 So, very probably, not (not M and C). (MT)

C3 So, very probably, M or not C. (DeM)

P3 And assuming we're right about C,

C4 Very probably M, or something close to it.

Here we clearly see how confirmation of such an unexpected prediction increases our state of knowledge of the objects and processes of the perceived world, as well as providing strong evidence for the existence of particular 'underlying' mechanisms and structures. Now we know about the crater (about the existence of the planet Neptune, about the holes in the ozone layer or whatever); about many things that were once no more than predictions. In this way, deductive logical reasoning makes a central contribution to the generation of new knowledge of the world. But notice also that it is not the reasoning alone that has increased our knowledge, it is the test: the physical contact with the new 'situation', entity or process, interpreted in the light of such reasoning, that is crucial.

Fallacies of Type 2 theory testing

Here we can follow Giere, once again, in identifying a number of fallacies of Type 2 scientific theory testing, that is, a number of 'recog-

nisable patterns' of reasoning that 'seem superficially to be alright, but do not in fact provide adequate support for the stated conclusion'.[3] As Giere says, 'the general mistake in these fallacious patterns is the failure to satisfy the second condition of a good test (a prediction not likely to be fulfilled if the theory is false). But the failure gets disguised in various ways'.[4]

We first consider probably the most familiar of such fallacies, those of *vague predictions* and *multiple predictions*. They apply particularly to claims for special future-predicting powers of psychics and astrologers. In the first case, the prediction might at first appear improbable if the psychic or astrologer lacks special powers, but in fact is sufficiently vague to stand a good chance of being fulfilled whether or not the theoretical hypothesis (that they do indeed have such special powers) is actually true. 'You will have some money problems, but things will work out alright. A new relationship is a serious possibility.'

With multiple predictions, the oracle makes quite specific and unlikely predictions, but makes so many of them (and/or repeats them for such long periods) that at least one is likely to be fulfilled. At least one of their selected famous people will (eventually) marry, die, be murdered or arrested in the course of one year, two years, three years. Then, of course, the oracle makes a great deal of the one or two predictions that are confirmed and keeps quiet about all the others that weren't. It looks as if an unexpected prediction has been confirmed. But this is not so. If we treat all their predictions on a given occasion as a single – conjunct – super-prediction it is clearly refuted. If we treat it as a disjunct – x or y or z or a or b, then it not unlikely that one of the disjuncts will be fulfilled and condition two is not satisfied.

Giere also identifies a fallacy of *no predictions*, and a version of *false dichotomy*, common in relation to theory testing. In the former sort of case, the theory proponent focuses upon some apparently inexplicable events – typically in the remote past – and then proposes a possible explanation. In fact, no new predictions can be derived from the theory, but it looks as if it satisfies condition two by virtue of the (alleged) mystery surrounding the past events in question; they could not have happened if the theory were not true. For example:

P1 The pyramids could exist only if aliens built them.

P2 The pyramids do exist.

C So aliens built them.

But P1 is simply false; there are other possible (and much more plausible, independently testable) explanations. So really they give us the fallacy of *affirming the consequent*.

P1 If aliens (of our favoured kind) had come to Earth in the past, then pyramids would exist, because they had built them.

P2 Pyramids do exist.

C So aliens came to Earth in the past, and built the pyramids.

In the latter sort of case, of false dichotomy, a theory proponent seeks to convince us of some radical explanation for a particular observation, for example, strange lights observed in the sky. They set up a series of possible explanations, including more 'straightforward' ones (weather balloons, clouds, aircraft, artificial satellites), dismiss all of the other ones and claim the truth must lie in their favoured 'non-straightforward' explanation (an alien spacecraft). Quite apart from the likely difficulty of any such definitive refutation of all straightforward possibilities, the problem is that there will always be further possibilities (in particular, some little known but still terrestrial phenomenon) which have not been considered. Independent verification of the theory is required.

Finally, there is the *ad hoc rescue*. The prediction is refuted and a new theory is concocted to explain why this happened, despite the (supposed) truth of the original theory. As we have seen, this, in itself, is not a fallacy, but rather an important part of scientific advance. But if the new theory is not independently testable, worse still, if the observations that refuted the original prediction are then presented as evidence confirming the modified theory, then a fallacy is involved. Again, this is an example of affirming the consequent.

P1 If the new (modified) hypothesis is true, then it follows that there will be observations that refute the original hypothesis.

P2 We have made an observation that refutes the original hypothesis.

C Therefore, the new (modified) hypothesis is true.

Creationism versus Darwinian theory

Many of the issues discussed so far in relation to Type 2 theory can be illustrated by reference to the contrasting models of Darwinian evolutionary theory and creationism, both aiming to explain the complexity, diversity, interdependence and functional specialisation of

living things and their constituent organs and structures. As Richard Norman says:

> Creationism is the view that the creation story in the first chapter of the Bible is literally true, that the whole universe, including our own earth with its species of plants and animals and the first human beings, was created by the direct agency of God in a period of six days, a few thousand years ago.[5]

According to this conception, God has created humans as creatures quite different from animals. Whereas animals are mere bodies (or perhaps bodies with feelings), humans are also, and essentially, immaterial souls, capable of intelligence, free will, moral judgment and a life after the death of the body.

So do creationists generally assume a fixity of the 'essential natures' or 'forms' of animal and plant species: fixed limits within which the descendents from common parents always remain. And they assume that everything in the world, including individual animals and species (as well as the organs of individual animals), has some aim or purpose within the grand scheme of things, typically the purpose of satisfying some human or divine requirement.[6]

The analogical basis of this theory is nicely brought out by Norman's succinct formulation of the design argument of the creationists:

> P1 Living things [and their component parts] and other features of the natural world are organised in such a way that they serve a purpose [eyes for seeing, feet for walking, trees for providing oxygen].
>
> P2 Where [the components of] human artefacts are organised in such a way that they serve a purpose, this is because they have been created by an intelligent designer.
>
> C Therefore those features of the natural world that are organised to serve a purpose must have been created by an intelligent designer.[7]

As Norman says, the analogy does have some appeal. But so too does it have major weaknesses.

> The trouble is that we have no idea how to fill in the details of the explanation. We cannot specify any of the physical processes, comparable to the carpenter's cutting and shaping of the wood, or the builder's assembling of the bricks and mortar. However powerful we may suppose the divine creator to be, we have no idea what physical techniques he might have used.[8]

According to the Darwinian model, by contrast, the world and the first life forms were created by purely 'material' or 'natural' processes in the very remote past, and have, since then, been slowly changing in various ways (through similarly material processes). Every group of organisms is descended from a common ancestor, including humans who share a (recent but now extinct) ancestor with the great apes. Rather than being 'fixed', species multiply, give birth to new species by splitting into daughter species, by budding, or simply by changing sufficiently through time.

This process is illustrated in the first of Darwin's two major analogical arguments, which compares species to branches of a single 'tree of life'. As Mayr notes:

> already in the summer of 1837 Darwin clearly stated that 'organised beings represent an irregularly branched tree' (Notebook B.21) and he drew several tree diagrams in which he distinguished living and extinct species by different symbols.[9]

Darwin's tree starts out as a single trunk some time in the remote past, thereafter generating many new branches. Some such branches continue to grow with little change for thousands or millions of years, some divide into two or more new branches, some change their form gradually over the millennia, and some make it through to the present day while others terminated long ago. Each branch is an individual species, or breeding population, at least until it splits or becomes so transformed in the course of time that descendents could no longer interbreed with their own (long dead) ancestors (the latter now being called 'speciation by anagenesis', or the accumulation of small changes). And a major cause of such splitting is geographic separation, leading to an eventual lack of compatibility between the populations concerned (called 'speciation through cladogenesis').

But most important was Darwin's account of the (underlying) mechanism that makes such species formation possible: the process of 'natural selection'. Again, Norman provides a succinct analysis of the analogical reasoning involved in this case.

> P1 We know that artificial selection of domesticated plants and animals can produce new varieties [with favoured random variations chosen for breeding].
>
> P2 We know that an analogous process of natural selection takes place in nature to produce new varieties better adapted to their environment.

P3 There is no reason why the process of natural selection that produces new varieties may not, over sufficiently long periods of time, also produce varieties so different as to constitute new species.

C The mechanism of natural selection can therefore explain how new species have come into existence with features adapted to their environment.[10]

Self-replicating systems tend to expand their populations at an exponential rate. But in practice, such growth is self-limiting as population pressure upon scarce resources produces intense competition, famine and disease. Darwin was well aware of the high fertility of animal and plant species and limitation of resources throughout the natural world, leading to fierce competitive struggle for existence within and between species.

Organisms in a population differ in their abilities to survive and reproduce (in the given circumstances). Given such intense competition for limited resources, even small differences can determine survival to reproductive age, or failure to do so. So are such differences, to some degree, passed on across the generations. And natural selection leads to increasing adaptation of populations to the demands of their environments. Such populations come to consist of individuals with properties that help them to survive and reproduce. And it is process of adaptation that gives the appearance of 'design'.

In stark contrast to the creationist model, which leaves the details of the creation process completely mysterious, the Darwinian model

> works because it invokes familiar processes, of biological reproduction and inheritance, natural variation, and the struggle for survival, and it shows how, given a sufficient time span, these mechanisms can account for the emergence of species adapted to their environment and possessing physical organs adapted to their functions. And the experimental biosciences, including modern genetics, can fill out, in immense detail, the picture of how these mechanisms work.[11]

Most important, Darwinian evolutionary theory has sustained a substantial record of generation of new knowledge through verification of clear, precise and otherwise unexpected predictions. The broad idea of generation of increasing complexity and diversity from greater simplicity and homogeneity has been amply confirmed by the fossil record going back billions of years. Many specific predictions of the nature of intermediate forms to fill earlier gaps in that record have been spec-

tacularly verified. And more recent genetic studies have provided a solid biochemical foundation for the theory.

To take some obvious examples, given fossil evidence of fish preceding amphibians, amphibians preceding reptiles, reptiles preceding mammals and birds, and apes preceding humans, evolutionary theory clearly implied the existence of a succession of intermediate forms in each case, their character more or less predictable on the basis of structural similarities and differences between previously known forms. And indeed, since Darwin's day, examples of such intermediate forms have been discovered, in increasing numbers, refuting the creationist idea of fixed, 'essential natures' of species, and confirming the evolutionary idea of the emergence of new species through natural variation and selection.

Some of the most spectacular recent cases here include 3.5–2.8 million-year-old hominid fossils of Australopithecus afarensis from Africa (a creature with a bipedal skeleton similar to that of modern humans, but a brain little larger than that of a modern chimpanzee) and the increasing numbers of feathered dinosaurs and dino-birds, clearly showing the dinosaurian ancestry of modern birds.

By contrast, creationist theory has provided no predictions, and hence no confirmed predictions. Certainly, the fossil record yields no radical and unexpected evidence of divine intervention. Instead creationists have responded to the increasing mass of new knowledge generated by the evolutionary model with the fallacy of 'ad hoc' rescue, claiming that the divine creator did, in fact, also put all these strange fossils into the ground as well, for mysterious reasons, with no possibility of independent testing of any such claims (or no empirical support for them whatsoever).

Theories in court

Until relatively recently, the criterion used to evaluate and determine the admissibility of expert scientific testimony in US courts was the so-called Frye Rule. This required that the 'major premise' (the scientific technique or theory used) must have 'gained general acceptance' and not be merely 'experimental' in nature. In *Frye v United States* (1923), the Court of Appeals for the District of Columbia noted:

> Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognised, and while the courts will go a long way in admit-

> ting expert testimony deduced from a well-recognised scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.[12]

As Walton points out, this principle was applied for more than 50 years, until the late seventies when 'courts began to repudiate or seriously question it and loosen the requirements for expert opinion testimony'.[13] And in 1975 new Federal Rules of Evidence were introduced that some saw as rejecting the Frye Rule. They mainly focused upon three requirements: appropriate specialised knowledge based upon 'experience, training or education' allowing an expert to 'assist the trier of fact'; this expert's obligation to disclose the reasoning behind their factual opinions on cross examination; the requirement for juries, rather than experts, to decide 'ultimate issues' in criminal cases.

> In the case of *Daubert v Merrill Dow Pharmaceuticals Inc (*1993) the US Supreme Court definitively rejected the Frye Rule, instead advocating as the criterion of admissibility of scientific testimony that it be 'not only relevant but reliable. According to the majority opinion, 'reliable' means 'derived by the scientific method' and 'supported by appropriate validation' … The Court made it clear that the intent of this ruling was to make the judge the gatekeeper to determine the admissibility of scientific testimony in a case. Amongst the criteria given [to define appropriate validation] were (1) 'testability' of the theory or scientific technique, (2) peer review and publication, (3) known or potential rate of error, and (4) general acceptance, not as the exclusive criterion [as in the Frye Rule] but as part of the court's assessment of reliability.[14]

Among other things, this decision is interesting because it seems to require that judges apply precisely the sort of criteria of the 'reliability' of scientific theories considered in this chapter, specifically their record of successful experimental verification (balanced against their record of refutation).

It seems to be true that the Frye Rule was dangerously restrictive – given the fundamental theoretical disagreements found in many areas of science, the subsequent rejection or substantial modification of virtually all theories deemed thoroughly reliable in the past, and the fact that even the best new ideas will take time to be understood, accepted and properly developed amongst particular communities of researchers.

On the other hand, the new rule places the judge in a position of great responsibility as far as these considerations are concerned. And

it is interesting that evidence that Merrill Dow tried to exclude because of an alleged failure to meet the requirements of the Frye Rule, evidence that proved decisive in finding Merrill Dow liable, ultimately turned out to be unreliable. Given that this centred upon a single paper describing a single animal test, showing 2 out of 8 rabbits receiving high doses of a chemical similar to the one under consideration in the case with foetal abnormalities, while a large body of published data described many properly constructed experiments showing no significant relation between the actual chemical and birth defects claimed by the plaintiff to be caused by it, the question arises of how well the Supreme Court judges understood their own principle.

In Australia, the Uniform Evidence Act (s 79) allows opinion based on specialised knowledge deriving from a person's training, study or experience, leaving 'specialised knowledge' undefined. Under the common law it is accepted that expert opinion must derive from a 'field of expertise' but as Odgers points out, 'Australian law has never clearly resolved the test for a "field of expertise" … There are authorities which appear to adopt a test of "general acceptance" in the relevant scientific discipline, authorities which require a court to make an assessment of "reliability" and authorities which adopt both tests.'[15] The Australian Law Reform Commission considered that such matters should be left to the discretion of the court, but some recent High Court decisions show the influence of *Daubert* in interpreting s 79 of the Act in terms of expert testimony meeting a standard of 'evidentiary reliability and relevance' to be admissible. In particular, the recent cases of *HG v The Queen* (1999) 197 CLR 414 and *Velevski v The Queen* (2002) 76 ALJR 402 involve the High Court applying a test of reliability closer to the *Daubert* approach.

Discussion topics

1. What are Type 2 theories? Find some examples and clearly explain the analogies involved.
2. How are Type 2 theories tested? Include reference to deductive reasoning in scientific theory testing.
3. What are fallacies of Type 2 theory testing? Find some examples of your own.
4. Was the Frye Rule dangerously restrictive? Were the new US evidence rules an improvement? Find some relevant examples.

5 In what ways do *HG v The Queen* and *Velevski v The Queen* show Australian courts moving towards a Daubert approach? See Odgers et al below.

Additional resources

S. Odgers, E. Pedern and M. Kumar, *Companion to the Uniform Evidence Act*, Lawbook Co, Sydney, 2004, ch 8.

R. Freckelton, *The Trial of the Expert: A Study of Expert Evidence and Forensic Experts*, Oxford University Press, Melbourne, 1987, pp 152–65.

SECTION TWO

Ethics, social theory and law

Chapter 8

METAETHICS, RIGHTS AND EQUALITY*

Ethics and law

There is a widely accepted view that science is quite separate from ethics. Science deals with, and attempts to explain, what is, what was and what will be the case. It considers what exists in the world, what happens and why it happens, what is possible and what is not. It tells us how to do what we want to do. It doesn't tell us what we should do, or what we ought to do. That is the role of ethics.

This needs some qualification. Science tells us that doing X will bring about Y. So there is a sense in which it tells us we probably should do X if we want to bring about Y (assuming that we can do X without serious costs or negative consequences and that we have no more important goals, incompatible with Y). But this is a conditional kind of a 'should' or an 'ought'. Ethics is about what we should or ought to do per se, either because the action in question is intrinsically good or because it is a means to achieving a goal or purpose that in intrinsically good or valuable.

It remains to be seen whether there really is such a fundamental conceptual divide between science and ethics. To the extent that there are facts about what is good for humans, other life forms or the environment, it will presumably be science that establishes such facts. Ethical conclusions about what we ought to do would then seem to follow from such scientific facts. We will return to this question, of whether an 'ought' can be derived from an 'is', later in the chapter and the book.

For the moment we should not get bogged down in worrying about the difference between *ethics* and *morality*. Different writers use the terms in different ways, but it is the ideas that matter. Most would see morality as about deciding what is right or wrong, good or bad, how we should live our lives and treat others. Ethics is that branch of philosophy that tries to understand, explain, interpret and guide moral decision-making and action.

Some see morality as confined to considerations of private life and personal relations. But moral philosophers recognise no such restriction. There are moral decisions and beliefs involved in all areas of human life, including particularly political and economic life. It is important that there is some solid moral foundation for decision-making and action in these areas if we are to avoid oppression, exploitation and misery on a large scale.

Whether we are fully conscious of this or not, we all have some understanding of morality and we all make moral judgments, and are affected by the moral judgments of others, on a day-to-day basis. Just as we can draw upon centuries of accumulated knowledge of logic to help in understanding, criticising and developing arguments, and of science to help in understanding the natural and social world, so can we draw upon centuries of serious thinking about ethics to help in understanding, criticising and developing our moral decision-making.

It is true that in logic and science there is substantial (though by no means unanimous) agreement amongst serious students as to the appropriate methods and results of the discipline, while this is not the case in ethics. This could be because of the strength of human self-interest, obstructing recognition or acceptance of moral truths.

But whatever the reasons for this lack of consensus, it is important to see that this is not the same thing as a lack of progress. Over many years of deliberation and debate, ethicists have made substantial progress in mapping out the range of possible ethical ideas and approaches to make sense of moral decision-making and moral action.

As we will see, some (legal positivists) say that law has – and should have – nothing, essentially, to do with ethics. But most people would be shocked to think that there could not be, or should not be, some kind of moral foundation or moral justification for at least some laws and legal procedures.

Others (natural law theorists) see law and morality as essentially the same thing: law is or should be institutionalised morality. But, again, many would see this idea as going too far in the other direction;

there are moral issues that should be kept outside of legal regulation and laws can be legitimately developed and applied in some areas without necessary ethical foundation.

Some might say that law defines and enforces minimum moral requirements that we all need to fulfil. But we should always aspire to go beyond such a bare minimum. Others would see some laws as immoral and unworthy of respect.

Probably, most people, thinking about the issue, would conclude that some core elements of law are or should be involved in protecting or enforcing certain fundamental rights and obligations and duties, or furthering the general good of the community, where legal regulation is appropriate to this purpose, and legal intervention might be expected to do more good than harm.

But the idea that law should aim to protect or enforce basic human rights, is quite compatible with the idea that it currently does no such thing. Perhaps, as Thrasymachus argues of all conventional justice in Plato's *Republic*, law serves merely to delude and oppress the majority so as to further the selfish interests of a powerful minority. Reference to protection of rights also highlights the issues of whose rights are being protected and at whose expense. And we need to consider the possibility that law does indeed institutionalise the moral values of some groups, but not of others. In this chapter we lay the foundations for considering and assessing all of these ideas, prior to considering them further in later chapters.

Metaethics

Philosophers have traditionally distinguished three major divisions within the study of ethics: *metaethics, normative ethics* and *applied ethics*. Metaethics is closely bound up with those areas of philosophy called metaphysics and epistemology. Metaphysics or ontology enquires into the fundamental nature of things, epistemology into the possible or appropriate techniques available for producing or establishing reliable knowledge of the world. Metaethics enquires into the fundamental nature of moral beliefs, judgments and actions, and into the appropriate means available for discovering or establishing the truth or appropriateness of, moral ideas, rules and values.

General normative ethics looks at particular ethical theories, setting out particular rules or principles for making and justifying moral judgments. These include theories of *act and rule utilitarianism*,

various so-called *deontological theories* of Kant and others, Aristotle's *virtue ethics* and more recent *feminist ethics* and *ethics of care*.

Applied ethics considers the application and development of such general theories in specific contexts of practical decision-making. This includes considerations of the sorts of moral issues that arise within the practice of different professions, of medicine, law, business and science, and the formulation of professional ethical codes and guidelines.

Often, the only ethical training provided in professional studies, including legal studies, is a brief consideration of professional codes at the end of three or four or more years of study. This is not just a case of too little too late. It makes no sense in terms of the logic of ethical studies, which derives, explains and justifies such specific applications in terms of broader and deeper ethical principles and practices of thought.

We will return to look at normative and applied ethics shortly, but here we consider metaethics, starting out with the fundamental nature of moral judgments, beliefs or ideas. As we all know, specifically *moral judgments* are distinguished by their reference to obligation (what we ought to do, or have a duty to do), to value (what is good or valuable) and to virtue (to the kind of person we ought to be). We have also noted that moral judgments are not the only sorts of judgments that make reference to duties and obligations and values. Science tells us what we ought to do to achieve particular sorts of ends. But generally science says nothing, or appears to say nothing, about precisely what sorts of – ultimate or final – ends we should pursue.

Judgments of what is right and wrong in law, of what we 'ought' to do in order to uphold the law, depend upon established legal principles. But as noted earlier, legal principles are not necessarily ethical principles. Some particular action or inaction (for example, not helping your terminally ill friend to end their life) might be the 'right' or 'correct' thing to do if you want to obey the law. But you might feel no moral obligation to obey the law because you judge it to be an immoral law.

As noted earlier, moral values are things of intrinsic worth, rather than things that derive value as a means to achieving other things. So-called *teleological* or *consequentialist* ethical theories see specifically moral or ethical obligation as derived from such ultimate values, considered as goals of action. Because some consequence of action is intrinsically good, perhaps the happiness, freedom or equality of all people, then we should necessarily and without qualification pursue it; we have a moral duty to do so. We must aim to maximise good consequences of action and minimise bad.[1]

Other so-called *deontological* theorists argue that some actions are, by their nature, intrinsically morally right or good, so that we ought (necessarily) to perform them independent of consideration of consequences. In particular, they argue that we have ethically fundamental duties, particularly duties to respect the rights and autonomy of others. Or they argue that it is the intention with which the act is carried out or the way in which consequences are brought about that is most important from an ethical perspective.[2]

Some argue that there is only one thing of intrinsic value; classical utilitarians argue that the only such intrinsically valuable thing is happiness, or pleasure and the absence of pain. Therefore all moral obligations concern individual responsibilities to try to maximise such happiness (for all creatures capable of experiencing it). Others recognise a range of different sorts of things as having intrinsic value, health, knowledge, empowerment, equality, loyalty, fidelity and kindness, which we are morally obliged to try to achieve.[3]

Virtually all ethicists agree that specifically ethical obligations integrally involve other people. They always go beyond merely or essentially self-interested considerations of our own wellbeing. It might be true that individual A ought to get more exercise in order to stay healthy. And good health might indeed be a moral value – something of intrinsic worth. But for a judgment about health to be a moral or an ethical, rather than a merely prudential, judgment, it must take account also of the health of others, or of the role of the health of the person concerned in realising other values that integrally involve the wellbeing, rights or autonomy of other people.

Many would go further and say that a condition called *universalisability* is also characteristic of all specifically moral obligations and judgments. This means that any moral obligation we attribute to ourselves or any other in any particular situation we must also attribute to any and all other people in the same type of situation. This is basically the same as the *Golden Rule*: do unto others as we would have them do unto us.[4]

Ethics and free will

As Kant argues, ethics is closely bound up with human free will: the capacity of people, not merely to respond directly to external and internal stimuli as causes with particular decisions and actions as effects, but rather to be able to deliberate about such stimuli as

reasons for actions, in the light of numerous different possibilities of decision and action; to adjudicate between competing or conflicting reasons for action (and the different actions they suggest) and formulate particular decisions or intentions prior to engaging in action, on the basis of such deliberation.

Desires arise spontaneously in people as a result of particular perceptual stimuli. We feel hungry and desire food. We might also believe that we can get food, easily and quickly, without other problems, by doing X. But the gap between reason and decision allows us to consider moral obligations or commitments that might conflict with or override the performance of X. They might more urgently command the resources necessary for accomplishing X: resources of time, effort, money or whatever.

This gap, as philosopher John Searle calls it, provides a space not just for adjudicating between conflicting reasons for action, but for rethinking the nature of moral obligation and commitment in the light of new information and experience. And such rethinking is not just a matter of internal thought processes, but also of discussion and debate with other people.

Just as the gap between desire and decision allows for the possibility of moral decision and action, so does the gap between decision and action (and gaps within ongoing sequences of actions) allow for failure of moral will or resolve, traditionally called *weakness of will*. Here, 'we judge unconditionally it would be better to do (the morally correct thing) y than x, believe that we are able to do either, and yet intentionally do x instead of y'.[5]

As Searle says:

> making up your mind is not enough; you still have to do it. It is in this gap between intention and action that we find the possibility, indeed the inevitability of at least some cases of weakness of will … As a result of deliberation we form an intention. But since at all times we have an indefinite range of choices open to us, when the moment comes to act on the intention several of the other choices may be attractive, or motivated on other grounds.[6]

Guilt and conscience

Ideas of morality are closely bound up in many people's minds with ideas of guilt and of conscience. The conscience is usually thought of as a particular internal mental agency or faculty with the power to

monitor an individual's feelings, thoughts and actions. Such an agency 'assesses' the moral status of such thoughts and actions, punishing wickedness with feelings of anxiety, guilt and self-loathing, and rewarding goodness and the renunciation of evil with feelings of pride and elevated self-worth. Supporters of conscience in this sense see it as a necessary counterweight to weakness of will: providing the moral strength to carry through moral intentions in action.

Some see such an agency as the voice of God, some as a biological product of evolution and some as a social creation, a policeman in the head, enforcing social norms and laws. Indeed, some identify the conscience-less individual with the psychopath or sociopath, ruthlessly exploiting and using others as means to their own immediate gratification, without reference to moral, social or legal rules.

Sociological theorists have highlighted the role of socio-economic factors in shaping dominant ideologies or belief systems within particular societies. Some have emphasised the importance of social anxiety, that is, of concern about the approval or disapproval of others (and avoiding experiences of shame), as a major element of conscience, motivating individual conformity to the dictates of such ideologies. Others have highlighted the power of dominant elites to shape and influence the conscience, and hence the behaviour, of the majority.

Sigmund Freud identified conscience with what he called the *superego*, a young child's image of parental values and parental discipline, 'introjected' or internalised as a way to try to come to terms with oedipal conflicts. Fearing parental retribution for their own jealous aggression and incestuous desires, the young child fantasises taking parental authority into their own mind to monitor and control their thoughts in such a way as to keep them safe from parental retribution in the external world.

Freud was highly ambivalent about the moral and social status of such an agency. On the one hand, he agreed that those with weak or absent superegos were likely to be prey to selfish instincts and highly anti-social. On the other, he saw too strong a superego as recalcitrant to advance beyond infantile moral beliefs and tending to overwhelm or preclude free action, initiative and fulfilling personal relationships with feelings of anxiety, guilt and depression.[7]

Other psychoanalytic traditions have been clearer in condemning the social superego as an agency just as likely to support immorality as morality. There is no guarantee that the values of parents or the dominant social values of the environment of the young child will be

genuine moral values, as opposed to destructive and corrupt values of racism, sexism, or fascism. A corrupt superego will leave individuals feeling guilty about thinking or doing the right (morally correct) thing, and feeling good about thinking or doing the wrong (immoral) thing. To the extent that they identify the voice of conscience with the voice of God, so will they see God as demanding support for such corrupt social practices and institutions.

At the same time, so have such traditions focused upon a minimal core of early parental love and care necessary for the infant's survival. The child's identification with the loving parent provides a – transhistorical and transcultural – basis for ethical feelings and actions, prior to all such social 'distortion'. It provides a foundation for further development of genuinely ethical ideas and actions.

Rights and freedoms

Ever since the 17th century, with the work of the great classical liberal theorists Thomas Hobbes and John Locke, the concept of *rights* has played a central role in considerations of moral obligations and values in the western philosophical tradition. In the liberal tradition, to say that person X has a right to Y, where Y is a specific action or inaction on the part of another person Z (or the product or result of such action or inaction), is to say that person X is – legitimately – entitled to get X, they can 'demand it as their due', and Z is under an obligation or duty to perform the action in question. Individuals may 'possess' rights merely by virtue of who they are (by virtue of being human or being a citizen, for example), or they may acquire rights by virtue of some action of their own and/or of others (by virtue of entering into contractual arrangements of some kind, or being promoted, for example).

Such legitimate claims upon the action of others may or may not be enforceable through the application of legal sanctions of one sort or another. Some have seen such legal enforcement as integral to the idea of rights, such that only enforceable legal rights are 'really existent rights'.[8]

The right of one person always puts another under a duty (not to infringe the right in question, or to do whatever is necessary to fulfil it). This duty can apply to specific people or to the world at large, as where an individual's right to protection of their life and property puts the rest of the world under an obligation to respect such life and property.[9]

Rights, and their corresponding duties, can be divided into the *posi-*

tive and the *negative*: the former being rights to another's positive action (as with the child's right to active care by their parents), the latter being rights to another's non-action (as with the child's right not to be abused by their parents). So can they be divided into *active* and *passive*: the former being rights to do something without interference, the latter being rights not to have particular things done to us. To the extent that such rights actually are respected – or legally safeguarded – in practice, the former are associated with so-called 'freedom to' perform particular actions (freedom of speech and association, for example), the latter with 'freedom from' interference (from becoming a victim of harassment, discrimination or negligent injury, for example).[10]

All freedoms 'to' involve corresponding freedoms 'from': from any restriction on the right 'to'. A freedom to walk the streets in safety is really also a freedom from attack or assault or harassment. All 'freedoms from' involve corresponding 'freedoms to': freedom to do whatever it is the thing you are 'free from' prevents you doing. Freedom from harassment allows you to walk the streets safely.

We can see how rights defined as active rights of some individuals could come into conflict with rights defined as passive rights of others, as a major source of moral – and political – conflict. A's active right to free speech might conflict with B's passive right to be free from racial vilification or endless oppressive advertising. A's active right to unlimited capital accumulation might conflict with B's passive right to be free from poverty and unemployment (as A takes over the business B works for and sacks B in the cause of rationalisation). One major aim of ethical theory is to find appropriately ethical, and legal, ways to resolve such conflicts.

A final important distinction to recognise in respect of rights, not generally emphasised by liberal theorists, is that between what we might call *actual* and 'merely' *formal* rights. Rights become 'actual' rights for particular individuals or groups when someone takes responsibility, or discharges a duty, to ensure provision, to those individuals or groups, of all material or other means necessary for the effective 'realisation' of such rights. For example, provision of a comprehensive and effective public health system, by the state, means realisation of everyone's right to quality health care, independent of what other resources they might have access to. Provision of a comprehensive and effective public education system means realisation of everyone's right to quality education, independent of their personal wealth or power.

On the other hand, individuals and groups may have legal rights to various sorts of things, in the sense of negative rights to others' non-interference in their exercise of such rights, while lacking effective access to the material means for the practical realisation of such rights, with no legal requirement for anyone to ensure such provision. Such legal rights therefore remain, for them, purely formal, rather than actual rights.

In a liberal capitalist democracy most people have the right to become cabinet ministers, CEOs of major corporations, owners of fancy sports cars and holiday villas. There is no fixed caste or apartheid system, for example, which lawfully precludes particular groups from access to such benefits. But it is quite clear that for the great majority of people these are purely formal, and in fact, quite empty and meaningless rights, because (a) such positions and luxury consumer goods are in very limited supply, with no possibility of increasing such supply within the present system (so it is impossible for more than a tiny minority to achieve such things) and (b) most people lack, and have no real hope of acquiring, the considerable 'extra' material and social resources needed to 'realise' such rights in practice.[11]

In such a political democracy all have a formal right to engage in some form of political communication, to seek to influence the political behaviour of others. But this is an actual right only for the tiny minority who control means of (mass) public communication that allow them to make contact with others on a large scale and a regular basis.

Rights for some people always entail duties on the part of other people. But there are often difficult moral and political questions of who should shoulder the burden of the duties in question. Those with different political persuasions might agree that everyone has a basic right to subsistence. But some right-wingers might see this as a purely negative right, such that no-one has a right to actively prevent anyone else from securing their own subsistence, but no-one (apart, perhaps, from the individual concerned) actually has any responsibility to actively ensure such subsistence (unless they have freely chosen to take on such a responsibility, such as other family members and private charities). By contrast, more socially oriented liberals will see the state, or the community, as having primary responsibility in this area, both in attempting to ensure full employment and in providing welfare to those unable to work.

In the absence of an adequate supply of reasonably paid jobs and comprehensive social security, a negative right to subsistence becomes a purely formal, empty right for some. And from a social-liberal perspective, the failure of a government to provide such things, when it is in a position to do so, could be seen as active prevention of some individuals from exercising such a right (that is, infringement of the negative right to subsistence – a right not to be abused by the state).

Law and rights

Prior to the 17th century, monarchs typically claimed a divine right to rule as source of all authority and all law, with the rest of the population subject to such absolute authority. In this context, and for most people, obligation was mainly about the duty of the subject to obey and conform to the dictates of such divine authority.

Following the English Revolution, the first of the great social revolutions of the modern age that overthrew such royal absolutism and shifted political power to representative parliamentary bodies, the focus of consideration of moral obligation increasingly shifted towards ideas of individual rights, freedoms and equality.

One key idea became that of a *social contract* between free and equal but insecure individuals, granting political authority to a sovereign to govern on their behalf as a means of rendering their lives more ordered and secure. Such free individuals have certain basic rights to life and justly acquired property, and rights to take action to protect such rights. They have these 'universal' rights simply by virtue of their humanity.

For Locke, the individual's primary right of ownership of their own body and bodily powers provided the basis for legitimate rights of ownership of external things. As the individual 'mixed their labour' with raw materials of various kinds, they effectively transferred their ownership of their own labour to that part of the external world that they had 'improved' with such labour. And with the right of ownership of such property came the right to take whatever steps were necessary to protect it from unwanted interference by others.

Subsequent theorists developed different theories of legitimation of private property ownership, based upon ideas of free contractual agreements, fair bargains, just deserts and utilitarian considerations (of general benefits from particular forms of ownership). We will consider these ideas in detail in later chapters. At this stage, we need merely note the liberal emphasis upon entitlement to others' non-

interference with assets legitimately acquired, leaving the owners free to benefit from and dispose of such assets as they see fit, with an entitlement to take appropriate action to prevent or rectify any such interference by others.

But without an established structure of authority, able to effectively enforce laws protecting such life and property, such basic human rights will always be under threat through unregulated competition for scarce resources. A minority of selfish individuals need the threat of punitive sanctions to be persuaded to respect the basic rights of others. The sovereign acquires their legitimacy by virtue of their role in securing the basic rights of all citizens, as laws, through effective exercise of a monopoly of armed force, with citizens retaining a right to remove or replace them if they fail in this.

The idea was to provide strong institutional support for such entitlements to 'non-interference' with private rights to life and property, so as to avoid the potential chaos of individual action in this area. And it is fair to say that in the western democracies in general, and common law countries in particular, such (negative) rights have remained the centre of concern for both the criminal and the civil law. Strong criminal sanctions are brought to bear upon at least certain sorts of 'intentional' or 'reckless' trespass to person (murders and manslaughters, assault, aggravated assault and sexual assault) and to property (theft, fraud, obtaining by deception, burglary, robbery). And the civil law provides compensation to victims of such trespass.

Liberal reformers have also supported voting rights for some or all citizens in periodic elections of representatives to legislative or executive, and sometimes judicial, bodies. They have supported the rights of some or all citizens to stand – and campaign – for political office and rights of free speech to allow for dissemination and debating of the different political programs.[12] And various legal steps have been taken to protect or enforce such rights.

Right-wing liberals, or libertarians, have been reluctant to extend legally sanctioned (universal) rights of all beyond this point, arguing that free markets are the most efficient way to accomplish all other aspects of social integration and co-ordination. In other words, all other rights are products of free negotiation between individuals, purchasing the benefit of others' duties or obligations through appropriate financial consideration. And, again, the civil law functions to enforce such financially based contractual arrangements for agreed transfers of property, and provision of services, and provide compen-

sation for negligent (and intentional) trespass to person or property or failure of such service provision.

Libertarians are also distinguished by their firm rejection of any idea of *special social responsibilities or duties on the part of those who derive special benefits by virtue of protection of such basic rights*, associated with or arising out of such special benefits (over and above general responsibilities of non-interference in the basic rights of others). For them, ownership of substantial private property carries no necessary responsibility to support public services available to others with much less. The extensive rights of free speech available to a newspaper publisher carry no responsibilities to ensure fair and balanced treatment of the issues of the day, with a diverse range of perspectives represented. Those occupying positions of political power owe no necessary duties to the powerless, to reduce inequality and poverty, expand social services and safeguard the natural environment. From a libertarian perspective, beyond a basic respect for others' life and property, individuals can only expect others to act responsibly towards them if they pay for the privilege, through contractual arrangements of some kind.

Left- or social-liberals have recognised a broader range of basic human rights, similarly calling for legal support. These include workers' (*positive*) rights to be able to work, to decent pay and healthy working conditions (with corresponding duties of employers to provide such things); workers' *(active*) rights to form unions, participate in collective bargaining and engage in strikes and pickets; consumers' rights to be properly informed of the nature of goods or services available to them, and protected from harm attributable to such goods and services (with corresponding duties on the part of producers and marketers); parents' rights to paid maternity and paternity leave from work; and citizens rights of access to public health, education and welfare services, proper legal representation and freedom from racial discrimination and vilification. Some have supported animals' rights to decent and respectful treatment.

For these rights to be effectively realised, those benefiting disproportionately from private property rights and market forces have to accept responsibilities toward those not doing so well. Thus, social liberals recognise the necessity for progressive taxation and regulation of markets in order to provide the public funding and economic stability to allow for enforcement of universal rights to employment and social services for those in need.

Many such 'moral' rights are identified in the *Universal Declaration of Human Rights*, accepted by the representatives of 48 states in the UN General Assembly in 1948. In addition to such classical liberal (*mainly negative*) rights as life, liberty and security of person, freedom from arbitrary deprivation of property, freedom of thought and expression (including religion), democratic government and universal suffrage, we find (*mainly positive*) rights to political asylum, social security, paid work, equal pay for equal work, trade union membership, rest and leisure, a satisfactory standard of living, free elementary education, participation in the cultural life of the community and access to a social and international order to realise these freedoms.

Also included is the right to *an effective remedy* in law to uphold rights, equality before the law and a fair trial. But not all signatories have taken steps to ensure any such effective legal remedies, particularly in the area of what are called 'economic, social and cultural rights'.

Some of Australia's treaty commitments to human rights have been enacted directly by legislation, but it is commonly accepted that 'unless specifically incorporated by a valid federal law, international rules (whether of treaties or of customary law) are not, as such, part of Australian domestic law'. In *Teoh v Minister of Immigration, Local Government and Ethnic Affairs* (1994) 121 ALR 436, the High Court held that the ratification of a convention, in this case the UN Convention on the Rights of the Child, was an adequate foundation for a legitimate expectation that administrators in government would act in conformity with the Convention. However, that case concerned only the right – as a matter of procedural fairness – to have the rights of the child under the Convention properly considered before the deportation of the child's parent. In the 2003 case of *Lam*, the High Court threw doubt on even that proposition.

Of course, this says nothing about the true moral basis of any particular right or duty. And we will return to this issue when we have considered some appropriate techniques of ethical assessment of particular claims in this area. For present purposes we note how crucial the concepts of rights and duties are within the common law system. Most areas of law can be seen as supporting the rights of particular individuals or groups through enforcing corresponding duties or obligations on the part of others. Some of these are general rights and duties of all citizens, some are special rights and duties associated with

particular relationships, including the relations of doctors and patients, lawyers and clients, householders and guests, employers and employees, shareholders and managers.

Equality and justice

This brings us to considerations of *equality*. As noted earlier, at a time when right liberals were pushing the concept of 'rights' to the centre of ethical concern with no corresponding social responsibilities beyond recognition of such rights of others, some left-wingers were concerned by the sort of issue raised earlier: the difference between actual and formal rights. They saw that without equal access to the material and social means for the realisation of basic rights, legal protection and general acceptance of such rights could become the means for enhancing and legitimating increasing social inequality and oppression. Hence their emphasis upon the importance of equality as well as freedom in the creation of a fair and just society.

Just as there are different concepts of rights so are there different concepts of equality. We will consider this issue in greater depth in later chapters. At this stage we can distinguish three fundamentally different conceptions of equality. First of all, there is *equality of outcome*. This implies that all receive the same amounts of important (available) benefits. Such benefits include economic and political power, consumer goods (including housing and consumer durables), health care and education, rewarding employment and leisure time.

This need not refer to effective (private) ownership and control of precisely similar 'bundles' of goods. It can refer to rights of 'equal' participation in political and economic decision-making and of access to public health and educational and employment resources. And the outcomes would include 'optimum' health and education and fulfilment for all concerned through equal fulfilment of the differing needs (health, education, housing and dietary needs) of different individuals.

This need not imply absolute restriction or absolute uniformity of ownership of consumer goods. Nor need it imply coercive enforcement of active participation in social decision-making. But to the extent that some seek to gain more than their fair share of 'core' benefits, at the expense of the rest, it does imply some form of active restriction of such selfish behaviour.

Such equality of outcome implies an equal division of detriments or sacrifices necessary to achieve such benefits, most obviously of the

unrewarding labour activities necessary in some cases. But again, there will be leeway for each contributing according to their specific abilities and limitations.

By comparison, *equality of reward* or desert – the second conception – implies that individuals are rewarded, with benefits of some kind, or punished, with detriments of some kinds, in proportion to their positive or negative social contribution. On the positive side, individuals could be rewarded for the extent of their effort and sacrifice in the service of society (equal reward for equal effort), or for the objective value of the actual results of their contribution (equal reward for equal results). On the negative side, individuals could be punished either for the extent of their negligence, recklessness or selfish wickedness or for the extent of the (objective) damage resulting from their actions.

Those with better outcomes than other people will always claim that they deserve such advantages. But, in practice, such inequalities of outcome are seldom the result of consistent application of any clear principle of just deserts.

As implied above, equality of access to material goods, to political power or to meaningful work provides *equality of opportunity* for consumption or enjoyment or effective use of such goods, for meaningful exercise of such political power or worthwhile productive contribution in the course of such employment. But this is not the idea of equality of opportunity generally encountered in the literature.

Equality of opportunity – the third conception – is usually taken to refer to *free access to competitive selection procedures for jobs or political offices* within a pre-established hierarchy, or at least, a developed division of labour. At one extreme, it merely refers to the absence of institutionalised discrimination, denying individuals access to such procedures or positions for reasons not relevant to effective fulfilment of the offices in question. At the other, it refers to institutionalised efforts to ensure that everything possible is done to allow all individuals to enter such procedures on equal terms. And this implies active steps to redress or compensate for prior disadvantage.

It is clear that equality in each of these senses is closely bound up with our ideas and intuitions concerning *justice* and *fairness*. We tend to equate justice (in different areas) of life, with different conceptions of equality; including equality in rights and access to the means of realising those rights.

Many would say that a just political system is one which accords

basic political rights *equally* to all, including freedom (or equality of opportunity) to exercise political decision-making power, and freedoms from political victimisation. Justice or fairness in work and in law are often associated with ideas of *just deserts*; of getting what you deserve in terms of rewards or punishments. But so too are they associated with basic, and equal, rights to strike, enjoy healthy working conditions and gain access to proper legal representation.

Many, influenced by contemporary neo-liberalism, would say that individuals should not receive benefits unless they deserve them. And without reward in proportion to contribution, such as is supposedly provided by market forces, some sort of centralised and ongoing political coercion will be necessary to ensure that everyone does their share of work and never takes more than their fair share of goods. Hence the importance of just deserts in the economic realm.

But it is important to see that other views are possible. Periodic or ongoing redistribution could indeed prevent inequality of outcome getting completely out of hand without draconian restriction of rights. More radically, real economic justice (and real political democracy) might involve equal access to all material productive resources, equal sharing of real decision-making power in relation to the use of such resources, and equal sharing of all good and all bad jobs. It might, in other words, involve an extension of economic rights parallel to the supposedly equal political rights of all citizens. And effective use of such resources might not, after all, require either differential rewards or central coercion if people come to equate their own interests with the common good of all.

Equality and the law

Supposedly everyone is equally subject to the same laws in all common law jurisdictions. Judges and magistrates swear oaths to 'do right by all manner of people … without fear or favour, affection or ill-will'.[13] And legal authorities have typically seen equal access to justice, including access to appropriate professional assistance (to allow all citizens to assert and defend their legal rights effectively in court), as an integral part of the 'rule of law'. But we need to look closely at what equality before the law means in practice.

There are no laws requiring ongoing redistribution in the service of equality of outcome.[14] Such legislation as does make reference to equality is typically concerned with equality of opportunity, in the

restricted sense of access to competitive selection procedures, and the operation of such procedures, rather than real equality of outcome. But some state and federal anti-discrimination legislation (relating to sex, age, race or ethnic origin, marital status, pregnancy, physical or mental disability and political conviction) also deals with direct access to goods or services.

The same punishment, for example, a fine of a particular sum of money, can have quite different practical – and hence ethical – significance when imposed upon a rich person as against a poor person; the former suffering not at all, the latter very much so. An illusion of equality masks radical inequality of outcome.[15]

Assuming that a desperate poor person and a comfortable rich person are 'equally' criminally responsible for stealing the same sum of money fails to take account of real ethical measures of 'equal' culpability. Action motivated by need, by ignorance or lack of available alternatives, is different from action motivated by greed and freely chosen despite full awareness of real non-criminal alternatives. Action that is not only criminal but also betrays the special trust associated with social positions of special power and privilege would seem to carry greater moral culpability.

The law has supposedly come to recognise the importance of market regulation, alongside of core criminal legal protection of life and property, with regulations dealing with misleading advertising, anti-competitive behaviour, occupational health and safety and environmental protection, consumer protection and fair trading. But even though there is substantial evidence that the proscribed behaviours cause greater harm than 'traditional' offences, these developments still remain outside of the 'moral' mainstream of criminal law, with lesser penalties, and fewer prosecutions.

As Glasbeek explains:

> by their very nature, [these newer regulations] are particularistic. They address specific problems arising out of discrete forms of productive activities … They merely provide a framework in which market activities can be conducted … The underlying premise [remains] that market capitalism is per se beneficial and should be promoted whenever possible … This understanding of market regulatory laws … legitimates special treatment for those regulated by such schemes. If fierce competition and aggressive behaviour is welcomed and promoted by the regulatory apparatus, there are bound to be a few people [or corporations] who use excessive aggression, who breach the regulations. This approach creates an atmosphere in which violations are unlikely

> to be seen as serious misconduct. So difficult is it to tell the difference between acceptable and admired aggression and unacceptable and sanctionable aggression that it will be hard for violators to admit that they have done anything wrong.[16]

Similar issues arise in relation to the consistent application of the basic categories of 'conventional' criminal law. It seems clear, in particular, that the idea of 'recklessness', as sufficient mens rea for murder (or assault) does not generally 'carry over' in consistent fashion from 'street' crimes to major harms resulting from profit-seeking business operations.

In this case Glasbeek considers the case of a building contractor retained to build an 80-storey structure.

> The builder's intention is to construct a building according to the given specifications for a price that will yield a profit. The builder knows, from experience and available statistics, that for every ten storeys over fifty, one more worker will have an accident resulting in injury or death.[17]

The conventional idea is that the builder should not be held criminally responsible for any such workers' deaths because they lacked any malicious intent to harm such workers. But, as Glasbeek points out:

> If a person gets into a car blind drunk and kills someone on the road because the object in front of the car seemed to be a cardboard box, the law would hold the driver criminally responsible. It would do so even though the driver neither meant to hurt anyone nor had formed any intention of doing so. It would do so because the driver had been utterly reckless about the wellbeing of other … [But] did our … builder not exhibit the same kind of indifference?[18]

Everyone does not have equal access to legal resources. Some (members of the ruling class) can afford teams of skilled barristers, detectives and forensic scientists gathering and processing evidence, crowds of expert witnesses testifying for them, and access to the different quality of justice offered in the higher courts. Others (members of the working class) are confined to brief chats with court-appointed solicitors and the summary justice of Magistrates Courts, without benefit of judge or jury.

To the extent that the law exists, essentially, to protect private property, its resources are disproportionately mobilised to protect the substantial property of the wealthy few, rather than the meagre property of the many poor. And it is law that defines property ownership

(in terms of groups of rights of disposal, benefit and exclusion), identifying shareholders as owners of public corporations, for example, when it could conceivably identify the workers or local community members as owners of such corporations, with rights of access to profits and participation in selection of board members and determination of company policy.

The nature of moral ideas

Ethical judgments and beliefs are distinguished by their reference to intrinsic values, to obligations and virtues, rights and duties, equality and justice. But there are other, equally important questions we can ask about *the nature of ethical judgments and beliefs*. Are they claims to truth? Are they expressions of feelings or attitudes, rather than truth claims? Are they prescriptions or imperatives, calls for others to think or act in particular ways?

Cognitivists generally see moral judgments as truth claims – supposedly true propositions – whose truth can be determined through the use of our cognitive faculties – of reason and perception. *Non-cognitivists*, on the other hand, see moral judgments as expressions of emotion, of our (positive or negative) *feelings* about particular issues, expressions of our attitudes, choices or decisions, seeking to evoke similar attitudes, choices or decisions in others, or universal *prescriptions*, through which we prescribe or require particular courses of conduct of specific others and of anyone else in similar circumstances.[19]

There are problems with all these ideas. To see moral judgments as mere expressions of individual feeling seems to undermine all possibility of serious moral deliberation, debate, decision-making or progress. But to see moral judgments as truth claims also raises difficulties. Most philosophers today see factual statements as referring to objects in a world that exists independently of our language and judgment, and what we say about such objects as true or false depending on whether it *corresponds* to how things are in the world. But then the issue arises of what particular things (or states of affairs) in the world render moral statements true or false? What states of affairs do true moral statements correspond to? And if such moral facts are out there to be perceived or understood then why is there not greater consensus amongst people about moral truths?

Plato argued that the facts in question are not to be found in this

earthly world, but rather in a heavenly world properly seen only after death. In this other world, such things as goodness, justice and beauty exist in their pure forms. Or rather, such pure forms, unlike physical objects, have a non-spatio-temporal existence, and are accessible only through the exercise of pure reason rather than perception. Such reason can show us the truth or falsity of moral claims, by reference to such pure forms. A particular earthly thing or action is good to the extent that it partakes of, reflects or approximates to the pure form of goodness.[20]

If we reject this two-worlds approach then we have to find the truth of moral statements in the ordinary physical world. And there are numerous different suggestions about how we might do so.

Moral psychology

As well as inquiring into the nature of moral judgments, metaethics considers issues of the origin, rational justification and truth conditions of such judgments. Clearly, origin, rational justification and truth conditions are closely related. If we assume that moral judgments are rendered true by virtue of their corresponding to moral facts, then the question becomes one of how we gain access to such moral facts.

There are three sorts of answers that have been offered to the question: *God* reveals such facts to us; our own *human nature* does so, through inbuilt faculties of feeling, intuition and/or logical reasoning; or *society* does so, either through some universal process of socialisation, or through some specific social-historical dynamic of development, creating the right social circumstances to allow recognition of fundamental moral truths.

These different forces need not be mutually exclusive. It is possible, for example, that our evolutionary history has equipped us with capacities or tendencies to recognise fundamental moral truths, but only certain sorts of social situation allow such faculties to properly develop, while others thwart or restrict their development in some or all people.

As we will see, a number of different approaches tend to emphasise either feelings or rational thought as fundamental sources of moral insight. But it is quite possible to see that both feeling and logical reasoning are equally necessary and interdependent in this regard. As 18th-century philosopher David Hume, among others, argues, our

ethical ideas could have their origins in particular sorts of feelings, perhaps feelings of sympathy for others or humanity or fellow feeling, as intrinsic parts of our human nature:

> Through sympathy we identify with and are moved by the happiness and suffering, joy and sorrow, pleasures and pains of others. Under the force of this feeling we tend to express our approval of those actions which promote the happiness of others and our disapproval of actions which have the opposite effect.[21]

But then:

> our judgment can … 'correct' our emotions, so that we recognise a person or an action to be admirable or deplorable, whatever the feelings we may happen to have … Our judgments can transcend our feelings and can have an impersonal character.[22]

Richard Norman argues that, in fact, 'there is a great range of such 'primitive responses [in addition to sympathy or empathy] … which underlie our shared vocabulary of evaluation', including respect for other people:

> Whereas sympathy is primarily a response to others as passive experiencers, as beings who are effected by a world in various ways and who enjoy or suffer accordingly, respect is a response to others as active beings, as agents. So, whereas sympathy involves a spontaneous inclination to respond to other people's needs and interests as our own, respect may in contrast involve distancing oneself and recognising that others' projects are theirs, not mine. It is the inclination, not to live others' lives for them, but to stand aside and let them live their own life in their own way.[23]

Others have traced what they believe are fixed psychological stages of moral development corresponding to different stages of cognitive development, through which all children and adolescents are supposed to pass. The most influential scheme is that of psychologist Lawrence Kohlberg. According to Kohlberg, for very young children, 'bad' means 'what incurs punishment', 'good' means obedience to parental rules. Later, 'good' comes to mean 'what gets you what you want' in the form of reward. Then it comes to mean 'what parents approve of', and then 'what is approved of in the wider society'. Finally, in adolescence, the individual moves on to rule utilitarianism and Golden Rule consistency.

There are problems with any such mechanistic models. In particular, it is far from obvious that children's moral awareness necessarily

develops in tandem with their cognitive abilities. Very bright, knowledgeable people can be selfish, ruthless and cruel.

Divine commands

For divine command theorists, moral standards are simply the commands of God that instruct humans as to how they should behave. And to test the truth of moral claims we have merely to consult the relevant religious authorities or sacred texts.

As Waluchow shows, there are, in fact, two different versions of divine command theory:

> The *moral ground* version … says that the grounds for our moral claims lie in God's will or commandments. If X is morally right, this is because God has commanded that we do X...

The *moral index* theory, by contrast, says that while God does indeed always command us to do the morally right thing, this is because that thing is independently morally correct.

> Being a supremely perfect being with unlimited knowledge and benevolence, God knows the true standards of morality and lets us know what they are.[24]

There are major problems with both theories. The moral ground theory seems to render moral values completely arbitrary. God could command us to torture and kill. And we cannot say that he would never do such things because according to this theory there are no independent moral standards with which to make such judgments. 'Good' just means 'what God commands' so if God commands torture then torture is good by definition.

As Waluchow points out, it would be good if we were sure, as in the moral index theory, that God always commanded us to the morally correct thing, even if we did not know why it was good. But how could we be sure that God commands the good if we had no independent means for assessing goodness? And how do we anyway know what God commands?[25]

> Where are we to discover the commands of God? In the Bible? In the Koran? How about the various sacred Hindu texts, or the disputed gospels of St Peter? Perhaps God's word is found in the pronouncements of divinely inspired prophets and religious leaders? But which prophets? Which religious leaders?[26]

Even if we – somehow – identify the true word of God, such a true word has still to be interpreted and applied in practice. The Bible tells us not to kill. But what about self-defence? And state-sponsored execution? What about just wars?

> [We] cannot entertain such questions without engaging in moral reflection ie some form of moral reasoning which is grounded in something other than the commands of God whose meaning is in question. So even if we believe that God's word is a guide to the requirements of morality, it cannot be a sufficient guide.[27]

It is easy to see the dangers of grounding moral belief, or legislation, in (supposed) divine commands. Whatever religious authority claims privileged access to such commands has no requirement to provide rational justification for whatever it claims to be the 'true' content of such commands. To the extent that others accept such an idea, they require no such justification and, in effect, hand over total and arbitrary power to the group in question. And any who challenge such alleged commands, instead of being answered with logical argument, are likely to be answered with repression and persecution.

Moral relativism

Moral relativists argue that moral principles and rules are basically sets of conventions, assumptions and practices accepted as right and appropriate within particular societies or communities at particular times. 'Correct' moral precepts can therefore be discovered by empirical observation of behaviour of the members of any society we happen to be in. Within that society it's right to do what the majority think it's right to do.

Some relativists see moral standards as *cultural conventions* internalised (to some extent) by all or most individuals growing up in a particular community. Such values are (typically) handed down to the new generation by parents and other authority figures. Others see such standards as *personal choices or commitments* of particular individuals selected from an available stock of cultural options. Once chosen, such values are 'right for' the person doing the choosing.[28]

Upon first consideration, relativism has a lot going for it. The great diversity of different values found across different cultures, suggests the absence of any sort of culture-independent, 'objective', rational foundation for morality. Moral relativism appears to be the only alternative to moral imperialism, judging our own values to be correct and

all others wrong, perhaps asserting our right and responsibility to impose our correct values on others by force. Certainly, it is difficult to see how the values of one society can be demonstrated to be true to members of a different society, given the absence of agreed criteria for assessing moral values. And as existentialists say, if God is the source of objective moral values, the end of belief in God is also the end of belief in such objective standards.[29]

On the other hand, to see moral values as no more than cultural conventions seems to undermine the possibility of moral criticism within and across cultures. What the people over there, or the majority of the people over here do or believe (or at least what they think of as moral values in which they should believe) can never be 'wrong', no matter how cruel or apparently irrational. It is morally correct by definition. And reference to 'the majority' highlights the fact that no society is actually homogenous in moral beliefs and actions, though there can be considerable homogeneity within particular social classes of class-divided societies (such as workers, peasants and capitalists). In complex and large-scale societies, in particular, there are hundreds of different subgroups, social strata and sub-communities within and between which people move, with quite different sets of beliefs, values and priorities. Which are supposed to be the correct values? And what about all those who claim to support particular values but fail to do so in practice?

If morals are just personal choices of cultural possibilities, there can be no real moral disagreements. We have opposite beliefs. But, by definition, we're both right! But how can this be?

It is crucially important here to distinguish the ways in which people acquire particular moral values from the truth of such values, although there will be close connections between the two. It is quite possible to see values as social products, with different values produced and transmitted in different social situations, without accepting that all such values are equally valid or that every society's 'dominant' values are 'true for that society', as relativists contend.

Some see moral values as influenced by particular social-historical circumstances, by individuals' positions within the social division of labour, and by historical forces of conflict, change and development. Marxists, in particular, see individuals' moral values as profoundly shaped and directed by social class positions, by ongoing inequalities and conflicts between different class groups, by changing economic circumstances and by new technological developments. In a class-divided

society, those enjoying the benefits of economic and political power and privilege will seek to justify and rationalise such power, both to themselves and to the exploited and powerless majority. Their economic power will enable them to propagate such ideas as 'ruling ideologies'.

Many would agree with relativists that parents and other authority figures can significantly influence the values of the growing child. And many would agree that at later stages of moral development individuals can free themselves to some extent from such influences to choose or develop systems of values of their own (or at least appropriate and develop available ideas in their own way).

But most theorists (including most Marxists) would still distinguish between the origin of moral beliefs and the truth status of such beliefs, with the latter having to do with what such beliefs say, rather than where they come from. Or they would distinguish between rational processes of construction of moral beliefs (perhaps building upon a foundation of empathy as suggested by Hume and some psychoanalytic theorists), making it more likely that they are in fact true, and less rational processes.

Some sorts of social situations will make it easier for individuals to perceive, discover or accept such objective moral truths, some will make it difficult or impossible. So, both true and false moral values are indeed social products, but nonetheless true or false for that.

Just because there are differences between the values of different culture groups does not mean that there are no correct answers. Some groups can be right and others wrong. And it is possible to stand up for what we believe to be right without necessarily being a cultural imperialist. Ideas can be changed through example, education, rational argument and dialogue.

And perhaps there is actually less disagreement in some areas than cultural relativists claim. A lot of supposedly moral disagreement turns out in practice to be factual disagreement, with different factual beliefs in different societies, or between different individuals, leading to different moral prescriptions and values. A and B both believe in the sanctity of human life but A opposes abortion altogether, believing that human life begins at conception, while B believes that the foetus is not truly a person in the first three months of its existence, and therefore does not necessarily oppose abortion during this period.

It is also possible that much apparent diversity in moral beliefs is really rather a sign of the necessity to adapt and apply the same basic moral principles to radically different natural and social circum-

stances, of war and peace, poverty and plenty, town and country, agrarian and industrial societies.

Relativism and law

The *Universal Declaration of Human Rights* of 1948 seems to offer the beginnings of a practical refutation of moral relativism, or at least some parts of it. For here we see agreement among the representatives of 48 nations upon a wide range of basic moral values and commitments. Of course, just because people of different cultures agree upon certain core values this does not mean that such values are actually correct, or that those concerned agree in the details of concrete implementation of such values. They may opt to subscribe to standards considered to be economically and politically feasible. But it could be that people do agree because they regard the propositions to be correct.

The development of international criminal law in the period following World War II, culminating in the agreement of representatives of 160 states meeting in Rome in 1998, by an overwhelming majority, to set up an International Criminal Court (operating from 2002), represents another arguable blow to ethical relativism. As Singer notes:

> The charter of the International Military Tribunal set up by the allies to try the leading Nazi war criminals at Nuremberg gave it jurisdiction over three kinds of crimes: crimes against peace, war crimes and crimes against humanity. In promulgating this charter, the Allies declared it a 'crime against peace' to initiate a war of aggression; a 'war crime' to murder, ill-treat or deport either civilians or prisoners of war; and a 'crime against humanity' to murder, exterminate, enslave or deport any civilian population, or to persecute them on political, racial or religious grounds. These acts, the charter of the tribunal stated, are crimes 'whether or not in violation of the domestic law of the country where perpetrated … ' Though the allies were able to draw on earlier precedents and conventions to justify their claim that crimes against humanity were already recognised in international law, the Nuremberg Tribunal gave new impetus to the idea that certain acts are so horrendous that they are crimes, no matter what the prevailing law at the time in the country in which they were perpetrated.[30]

Leading Nazis and their followers had been responsible for all these sorts of crimes, for which they had found both legal and 'moral' justification, creating serious problems for thoroughgoing moral relativists who would presumably have to say such activities were 'right for them' if not for their victims.

Of course, other actions that could have qualified as war crimes, such as the British fire-bombing of Dresden and Hamburg or the dropping of atomic bombs on Hiroshima and Nagasaki, went unpunished because the victors are the ones in a position to enforce laws and values upon the vanquished. Nevertheless, the Nuremberg principles were subsequently accepted by the member states of the United Nations, which asked the International Law Commission to formulate principles of international law relating to these and similar crimes. This commission recommended international criminal responsibility for crimes against humanity committed at the instigation or with the toleration of state authorities, meaning that any state has the right to try a person who has committed crimes against humanity. And the 1984 Convention Against Torture, signed by 110 states, accepted this principle,[31] though not all countries enacted legislation to this effect (recognising the principle of ''universal jurisdiction').

Singer argues:

> Like the Nuremberg tribunal, more recent international tribunals have arisen in the wake of tragic events: the wars that followed the break-up of the former Yugoslavia, the massacre in Rwanda, the Serbian attacks on the Albanian inhabitants of Kosovo, and the killings in East Timor by militia supported by the Indonesian armed forces. By strengthening the resolve of all decent people not to allow such tragedies to continue, these tribunals are pushing us towards a global system of criminal justice for such crimes. In contrast to the Nuremberg Tribunal, the trial of Slobodan Milosevic … is not justice exacted by the occupying forces against the leaders of a nation that has been forced into unconditional surrender. It is a sign of recognition, at least within Europe, that national sovereignty is no defence against a charge of crimes against humanity.[32]

The International Criminal Court (ICC), associated with the United Nations and situated in the Hague, marks the step beyond such one-time, ad hoc arrangements to the establishment of a permanent institutional base for prosecution of individuals accused of genocide,[33] crimes against humanity,[34] and war crimes,[35] on condition that such individuals are nationals of states that have ratified the treaty that establishes the ICC, or have committed crimes within such states or are referred to the court by the Security Council.

A crucial issue here concerns the circumstances in which it is recognised as legitimate, or mandatory, in international law for individual nations, groups of nations or the United Nations to intervene to stop

such crimes about to occur or already in progress. Singer argues that 'if punishment can be justified, so can intervention', and follows UN Secretary-General Kofi Annan in proposing that 'humanitarian intervention is justified when it is a response [with reasonable expectations of success] to acts that kill or inflict serious harm on large numbers of people, or deliberately inflict on them conditions of life calculated to bring about their physical destruction, and when the state nominally in charge is unable or unwilling to stop it'.[36]

In response to moral relativist suggestions that such intervention in other countries to protect human rights is a form of unacceptable cultural imperialism, Singer acknowledges that 'distinctive cultures embody ways of living that have been developed over countless generations, that when they are destroyed the accumulated wisdom they represent is lost and that we are all enriched by being able to observe and appreciate a diversity of cultures'. Moreover 'western culture has no monopoly on wisdom and still has much to learn'.

However, he also notes that 'once we accept that there is scope for rational argument in ethics, independent of any culture, we can also ask whether the values we are upholding are sound, defensible and justifiable'. And he concludes:

> Although reasonable people can disagree about many areas of ethics, and culture plays a role in these differences, sometimes what people claim to be a distinctive cultural practice really serves the interests of only a small minority of the population, rather than the people as a whole. Or perhaps it harms the interests of some without being beneficial to any, and has survived because it is associated with a religious doctrine or practice that is resistant to change. Acts of the kind carried out by Nazi Germany against Jews, Gypsies and homosexuals, by the Khmer Rouge against Cambodians they considered to be class enemies, by Hutus against Tutsis in Rwanda, and by cultures that practice female genital mutilation or forbid the education of women are not elements of a distinctive culture worth preserving and it is not imperialist to say that they lack the element of consideration for others that is required of any justifiable ethic.[37]

It is true that there remains good reason to suspect the selective enforcement of international law against regimes once they become obstacles to the geopolitical strategic goals of the major powers. This was arguably the case with Milosevic and Saddam Hussein, who were both allies of the West for much of their time in office.

While the ICC has been broadly supported by human rights organisations and lawyers' associations within the country, the Bush admin-

istration has refused to allow the United States to become a member, with Congress passing a law forbidding co-operation with it at any level of government, and another allowing Washington to cut of all military aid to countries not pledging to refuse to deliver US citizens for trial by the court. Rather than encouraging the development of a principled definition of the crime of terrorism, with the ICC then prosecuting suspected terrorists, the Bush administration has insisted upon establishing its own military tribunals to prosecute those it deems to be terrorists.

The stated justification for this policy is that the court will 'open American officials and military personnel in operations abroad to unjustified, frivolous, or politically motivated suits'.[38] But the real reason would seem to be that any such principled definition would apply directly to the activities of US officials and their allies overseas since at least 1945. Indeed, even without such a definition, as Gareau points out, the US campaign against the ICC 'is prima facie evidence that members of the administration believe that they themselves and some of their subordinates could be convicted in legitimate legal proceedings of war crimes, crimes against humanity, or genocide'.[39]

Far from supporting any principled moral relativism, this suggests that US officials are quite well aware of the 'objectively' immoral, and criminal, actions in which they and their supporters have been involved.

Discussion topics

1. What is ethics and how does it relate to law? Include reference to metaethics, free will, guilt and conscience.
2. What are rights and freedoms? What is equality? Could the protection of rights come into conflict with pursuit of equality? Include reference to ethics and law.
3. What kinds of things are moral ideas? Include reference to moral psychology, divine commands, moral relativism, natural law and social contract theory.

Additional resources

S. Mann, *Economics, Business Ethics and Law*, Lawbook Co, Sydney, 2003, chs 2 and 4.

P. Singer, *One World, The Ethics of Globalisation*, Text Publishing, Melbourne, 2002, ch 4.

Chapter 9

NORMATIVE ETHICS

Utilitarianism

As noted earlier, normative ethics is about the fundamental rules, principles and values underlying and justifying moral judgments. One of the most influential normative ethical theories is *utilitarianism*, originally developed in the 18th and 19th centuries through the work of Jeremy Bentham and John Stuart Mill.

As Waluchow points out, utilitarianism decisively rejects divine command theory, natural law and moral relativism, in favour of establishing what is seen as a straightforwardly empirical foundation for moral decision-making. This is the principle of *utility*, as universal objective standard for rational determination of moral rights, obligations and duties.[1]

Classical utilitarianism involves a monistic consequentialist theory of obligation, resting upon a particular theory of value. For classical utilitarians, the only intrinsic value is that of happiness or pleasure and the absence of pain and suffering. For Bentham it is only the quantity of such happiness that matters, while Mill considers also its quality, with 'higher' mental forms of happiness counting for more than 'lower' bodily forms. Some later utilitarian theorists have argued rather for pluralistic theories, which see a range of human experiences in addition to pleasure as also of intrinsic worth, including knowledge and aesthetic experience. Others see the satisfaction of rationally chosen and socially responsible individual preferences of all kinds as the basic goal of moral action. Others emphasise objective human wellbeing, in the form of health, welfare, empowerment and fulfilment as such an ultimate goal.

Act utilitarianism identifies an act as right and good 'if and only if there is no other action we could have done instead which either (a) would have produced a greater balance of utility over disutility, or (b) would have produced a smaller balance of disutility over utility'.[2] The second condition deals with situations where we cannot help but do some harm, and we are trying to minimise such harm.

Under *rule utilitarianism*, by contrast, it is not consequences of individual actions that matter: it is the consequences of everyone adopting a general rule under which the action falls. If not everyone could do what we propose to do now without great disutility resulting, then it's wrong for anyone to do it. For rule utilitarianists an act is morally right 'if and only if it conforms with a set of rules whose general observance would maximise utility'.[3]

But all utilitarians agree that the wellbeing of each and every sentient being affected by our actions should count equally in our decision-making. We should consider each and every person as equally deserving of happiness or fulfilment or wellbeing, and act accordingly.

In theory, both act and rule utilitarianism offer clear means for establishing moral priorities and resolving moral disputes. If we are a classical act utilitarian, we simply compare the overall happiness levels likely to result from each of the actions open to us and then perform the one which generates the greatest happiness.

But we do not have to look far to see the problems with this idea. First and most obvious is the problem of measuring, quantifying and comparing happiness levels of different experiences and different individuals. It is far from clear that this is possible. Even if it were, trying to calculate the precise long-term happiness consequences of numerous different possible actions could be fantastically difficult and time consuming. How could any decisions ever be made? And is setting out to maximise happiness really the best way to actually maximise happiness? Individuals typically achieve happiness, not by setting out to do so, but rather by doing specific sorts of useful things well. Why should things be any different for an individual trying to maximise the happiness of others?

Even if we could both quantify happiness and successfully plan to achieve it, the utilitarian emphasis upon maximisation could come into conflict with pursuit of equal shares of happiness for all. If action A produces one unit of happiness for Bill and six units for Ted, while action B produces three units for each of them, then action A should be chosen over action B upon strict principles of maximisation, but

action B should be chosen on grounds of equal rights to happiness.

Things are actually worse than this, since the act utilitarian maximisation principle would also seem to allow total sacrifice of the happiness, or objective well being or interests, of the few, to increase the happiness of the many. A hypothetical example frequently cited here is that of police authorities framing an innocent black man for murder of a white man to prevent violent race riots in a town in the southern US. Another example is Lord Denning's infamous statement about the need to preserve public confidence in the police even at the expense of allowing the wrongful conviction of the Guildford Four through fabricated police evidence.

Richard Norman observes:

> what utilitarianism fails to take on board … is that though some people's interests sometimes have to be sacrificed for the interests of others, there are limits. There are some things which, morally, you cannot do to people for the sake of the greater good.[4]

As Norman points out, this idea of 'the moral limits to the permissible treatment of human beings … can … be articulated in the language of human rights', as considered earlier.

> One significant use of the language of rights … is to set limits to utilitarian calculations. So to talk of each individual's right to life is to recognise that human beings are not just items to be weighed against one another in a utilitarian calculation of total net benefit. If I have a right to life, that means that other people may not deprive me of that right; my life is mine, not just a component of the general happiness, and it is for me to decide what is to done with it. Likewise the rights to certain basic freedoms establish the constraints on the ways in which human beings may be treated, and to violate those rights is to begin to rob them of their humanity.[5]

Act utilitarians fail to take account of special responsibilities created by past actions, of promising, for example, and by established relationships with family, friends or colleagues. An example here is that of the parent who has promised to buy their child a particular birthday present, but gives the money to a homeless person on the street instead. Another example might be that of saving a family member rather than a leading heart surgeon or wealthy philanthropist from a burning building. Perhaps the surgeon or philanthropist will save lives and produce masses of happiness or wellbeing, but your family members may legitimately expect you to rescue them first.

Another problem is that of free riders or law-breakers. An individual can reason that their theft of an item from a big store will greatly increase their happiness while contributing little or no sadness to any particular others, and is therefore not only legitimate but mandatory from a utilitarian perspective. Similarly, it would seem that utilitarians should steal from the rich to give to the poor if the rich cannot be persuaded to voluntarily redistribute their wealth to those in greatest need.

An act utilitarian response in these cases would be to refer to long-term consequences. By betraying the trust of family and friends through helping beggars, by colluding in framing the innocent, by stealing from stores or rich folk, we may well do more harm than good in the longer term. But such act utilitarians generally do accept the rightness of sacrificing some happiness – including some lives – to produce greater happiness, through saving more lives, for example. Military intervention in genocide situations, considered earlier, could involve some loss of life among the intervention force, to save many more lives on the ground. Shooting down the hijacked airliner could save many more lives in the urban centre for which it is headed.

As far as inequality is concerned, act utilitarians would remind us of the declining marginal utility of wealth: the fact that extra wealth for those who already have plenty will increase their happiness less than the same resources given to those in poverty and need. Similarly, they can refer to the epidemiological research of Richard Wilkinson and others showing that greater inequality of income leads to increasing disparities of health, with those at the lower end dying 5, 10 or 15 years earlier than those at the top, and significantly earlier than the poor in a more egalitarian system. So that far from increasing inequality, utilitarian judgments as guides to social policy will generally support redistribution in favour of material equality.

Rule utilitarianism appears to provide another possible response to some of the problems. We cannot condone general principles of robbery and 'fitting up' and betrayal; if everyone did these things this would lead to great social unhappiness. Rather, rule utilitarianism requires us to respect individual rights to life, property and justice.

Here we recognise also the impossibility of being able to predict precise long-term consequences of every act. Instead we settle for general rules that will generally, or in aggregate and in the long term, be more likely to do a bit more good than harm.

But rule utilitarianism has its own problems. By supporting some

rights it inevitably overrides others, and loses the benefits of principled resolution of conflicts of rights. If we insist on sticking with the rules no matter what special circumstances arise then we would seem to be following the rules for their own sake rather than for utility maximisation. But if we allow for deviation then we are back with act utilitarianism.

And we cannot follow ideal rules, which would create the happiest society if everyone followed them, given that not everyone is following them now. If private property were equally distributed, it might make sense to respect the property of others. Given that it is radically unequally distributed, taking from the rich to give to the poor might make moral sense.

We can see why many have recognised a close relationship between rule utilitarianism and law. In fact, some would say that the laws of the land are simply rules whose general observance and/or consistent legal enforcement maximises utility in the long term. Breaking particular laws on particular occasions or failing to enforce such laws may lead to an increase in happiness levels. But in the longer term, happiness will be maximised if all or most laws are obeyed by all or most people and/or consistently applied by the legal system.

Such claims, about the common good, have the benefit – at least in theory – of being open to empirical confirmation or refutation. As suggested earlier, this may be difficult in practice if we stick with the classical utilitarian emphasis upon happiness levels. But if we focus, instead, upon objective means of social welfare, it may not be so difficult after all.

When judges say they are making decisions on grounds of 'public policy', the most obvious ethical interpretation would be that they are aiming to establish general legal principles, the consistent application of which by other judges will maximise overall social welfare. When legislators recognise and support certain rights in law it could be that they believe that such rights are intrinsically valuable, or it could be that they believe that consistent legal recognition and support of such rights will maximise overall utility in the long term, even if ignoring such rights could increase utility in individual cases. Alternatively, legislators might support general legal principles that explicitly override particular individual rights in pursuit of the common good.

In theory, judges and legislators do not have the problem of failure of other legal practitioners to follow such principles because of the institutionalisation of consistent application (through stare decisis, etc). But there still remain difficult issues of discretion and interpretation that undermine such consistent application in practice. Most

importantly, it cannot be assumed that simply outlawing particular harmful behaviours or enforcing particular significant rights will necessarily maximise social welfare. The real social costs of such legal intervention could simply outweigh the social benefits, for many possible reasons.[6] This has been recognised in recent years particularly in relation to criminalisation of drug use, with an increasing (utilitarian) emphasis upon genuine harm minimisation.

Kant and Ross

Kant's deontological theory presents a radical alternative to utilitarianism. Whereas utilitarians accept no absolute and unconditional moral rules (apart from happiness maximisation), Kant supports absolute and exception-less principles; lying and committing suicide, for example, are never morally justified under any circumstances. Whereas only consequences matter for utilitarians, consequences are irrelevant for Kant. It is principles of action that matter: the way things are done. More specifically, actions have moral value by virtue of conforming to 'maxims' – general principles that specify (a) what we see ourselves doing and (b) our reasons for doing it.

Rationally based maxims must conform to what Kant calls the *Categorical Imperative*, a monistic theory of obligation. This requires universality and respect for the autonomy of other people. As Kant says: 'I ought never to act except in such a way I can also will that my maxim become a universal law.'[7] In other words, Kant thinks that we can discover moral rules through application of the purely rational, logical requirement of avoiding contradiction. When I think about lying, I must think of the consequences of everyone else lying in such a situation. And I will then find that in some sense I would contradict myself by lying. Lying presupposes people trusting each other to tell the truth. But with lying as a universal law there can be no such trust. I am relying upon an institution I am simultaneously destroying.

The Categorical Imperative requires us never to treat people as mere means to our own objectives. What makes them people is the fact that they have their own rational faculties to determine their own moral and practical ends. So, again, there is a kind of contradiction involved in treating people as mere means to our own ends. Instead, we must always respect their goals and purposes. To respect others as persons is to respect their worth and dignity as rational beings – beings able to rationally choose their own goals and purposes. We

must respect their autonomy, or power of rational self-governance.

There is general agreement among ethicists that universalisability is a necessary condition of valid moral judgment. We cannot morally act for a reason unless we accept that everyone else in the same situation should act in the same way. There are problems with this idea, nonetheless. In fact, it seems to come into conflict with respect for others' autonomy. The way we might want others to treat us might not be the way others want others to treat them. Perhaps we should rather treat others as they would want us to treat them.

Ethicists also generally agree about the central moral value of respect for autonomy. But different theorists have emphasised different senses of autonomy. Some focus primarily upon an absence of external coercion, allowing the individual to freely act as they choose to do. Others emphasise the provision of facilities that allow individuals to exercise a genuine freedom to choose between viable alternatives. And others again focus upon an absence of internal constraints that obstruct effective deliberation and selection of both the best means to chosen ends and of those ends themselves. We have already considered such possible 'internal' constraint in the form of an oppressive and infantile superego. Others would include stresses that interfere with the effective operation of our mental deliberative faculties, such as drugs, trauma or mental illness. Where other people, in a position to do so, fail to provide us with appropriate information, where they lie to us or otherwise deceive us, this too limits the effective exercise of the abilities required for rational deliberation.

The value of autonomy is recognised in law. In particular, requirements of informed consent to medical procedures, and consumer protection law, which requires reliable information to be provided to purchasers of goods, focus upon this last consideration of removing obvious obstacles to rational deliberation.

On the other hand, the law fails to adequately consider the way in which lack of autonomy, in terms of how the 'duress' of social situation or radical restriction of available options, may contribute to 'criminal' behaviour in a radically unequal and class-divided society. With its black-and-white picture of absolute 'free will' on the one hand, or complete absence of such free will on the other (in situations of duress or insanity), the criminal law radically fails to consider the ways in which real life options are more or less narrowly restricted for individuals and groups by virtue of their class positions.

At the other end of the social scale, so does the system fail to

address the encouragement to law-breaking offered by the effectively unrestricted scope of 'free' and autonomous action available to the rich and the powerful. They can come to take it for granted that they can 'do whatever they want'.

For WD Ross, both consequences and intentions are important in moral decision-making. But Ross' theory is particularly distinguished by its emphasis upon the centrality of special relationships in creating particular moral rights and duties.

Utilitarianism treats everyone as equally a beneficiary or victim of the action of everyone else. But in fact every person stands in different and specific moral relations to different people (and to the same ones). Particular sorts of relationships – of friend to friend, doctor to patient, teacher to student, colleague to colleague, child to parent, husband to wife, promisor to promisee – carry with them different and specific moral values, commitments and responsibilities, And such values and commitments can be important enough to override utilitarian calculations of maximisation of overall happiness, or objective social welfare.

Ross identifies six so-called prima facie duties, deeply grounded in our past actions and social relationships. These are duties of fidelity (telling the truth, keeping promises, honouring commitments), reparation (compensating for our wrongdoing), fair distribution of goods, beneficence (improving the condition of others), self-improvement and non-maleficence (not harming others). They are all things we should generally try to do. But different duties will be more significant, depending upon the nature of the situation. We should generally try to tell the truth and keep promises, for example, but where doing so will produce great harm to others we should prioritise non-maleficence.

Ross sees such duties, and the dominant duties in particular situations, as generally evident to our faculties of moral intuition. Others are not so sure and suggest that while these might be good rules of thumb, something like a utilitarian calculation is still necessary to decide what to do in difficult cases of conflict.

Aristotle and feminist ethics of care

Whereas Kant, Ross and the utilitarians all consider the question of what moral rules or principles should govern our choice of actions, Aristotle focused rather upon virtues or qualities of character that determine what kind of person we are. Rather than rules or principles to follow, he considered virtuous habits of life to develop or achieve and vices to try to avoid.

For Aristotle virtue is essentially a disposition or tendency acquired through practice and reason to choose well. We can relate this to John Searle's identification of the crucial gap in human conscious experience between desire and decision, noted earlier. Virtuous actions are not simply acts that conform to particular principles of value. By its nature moral action presupposes the shaping of feeling by rational deliberation: serious consideration of available alternatives. For Aristotle such deliberation is as important as the action itself. But all such deliberation is essentially concerned with choice of means to the end of a particular sort of wellbeing to which all are directed by their inherent human nature.

Aristotle's 'doctrine of the mean' sees virtue as consisting in observing the mean between excess and deficiency in our emotional responses to others and to the world. Some have understood this essentially as a doctrine of 'moderation in all things'. And Aristotle encourages such an interpretation by identifying a number of virtues as the mid-points between specific vices. One such virtue is temperance, for example, which concerns appetites for food, drink and sexual activity, standing midway between the excess of gluttony and the deficiency of starvation and frustration. However, Aristotle also says that observing the mean does not necessarily involve always choosing a mid-point, and that sometimes the right way to feel anger or fear or other emotions is to feel them very strongly.

Richard Norman seeks to reconcile these ideas by arguing that what Aristotle was really doing was rejecting Plato's earlier idea of reason and emotion as essentially antagonistic, with the requirement for reason to dominate and control emotion.

> For Aristotle feelings can themselves be the embodiment of reason. It is not just a matter of reason controlling and guiding the feelings. Rather, the feelings can themselves be more or less rational. Reason can be present in them … Essentially it is a matter of their being more or less appropriate to the situation. Take the case of anger. Suppose I become furious because someone fails to say hello to me … Here my anger is irrational. Suppose, in another case, that I become furious because I see a gang of children heartlessly taunting and bullying a younger child. Here my anger may be quite appropriate; the cause may be genuinely appalling.[8]

In other words, the doctrine of the mean commands that we avoid what would on the occasion in question be an excess or a deficiency, of tolerance, courage, benevolence, friendliness, generosity, or loyalty. Our feelings are rational 'in the sense that they are sensitive to the real

nature of the situation, not distorted by extraneous considerations'.[9] Discovering the mean, and responding accordingly, is a matter of what Aristotle calls 'practical wisdom' rather than theoretical knowledge. Whereas the latter is about discovering universal principles, the former is rather about sensitivity to the particularity of the situation. Such practical wisdom is a product of life experience. As Norman observes:

> The knowledge which enables us to understand [the doctrine of the mean] is acquired not by learning theoretical principles, but by moral training, by being properly brought up in a morally civilised community.[10]

In such a civilised community, the virtuous life, in this sense, is also the fulfilled and happy life.

The Aristotelian account has been criticised for failing to provide clear guidance in difficult situations of moral conflict and requiring supplementation with a theory of right action and obligation. But others have seen it as a strength of the theory that it recognises no clear or simple answers, since there are no such answers.

A number of feminist theorists have been sympathetic to Aristotle's rejection of abstract, universal principles and rules in favour of emotional responses as central features of ethical life. But in contrast to Aristotle's emphasis upon manly virtues of courage on the battlefield and temperance off it, they have tended to emphasise rather actions and values of 'communication, compromise, caring, special attention to concrete details and the various personal relationships and emotional bonds' involved in particular situations.[11]

Carol Gilligan critically re-examined Kohlberg's ideas of ethical developmental stages considered earlier. In particular, she compared the different responses of boys and girls to moral problem situations, including one involving a man who cannot afford to buy an expensive drug from a pharmacist to save his wife's life. She found that whereas boys in the later stages of moral development 'looked for answers in abstract principles governing property and the value of human lives' – in this case calling for the man to steal the drug – girls of a similar age looked for solutions through 'mediation, communication, compromise and the personal relationship' between the man and the pharmacist.[12]

> Under Kohlberg's model, [the girls] had not ascended to the more advanced stage of moral reasoning [attained by the boys]. According to Gilligan, [the girls'] response is in no way inferior or less advanced than [the boys]. [They] appealed to what many feminists now call an 'ethics of care'.[13]

The boys' appeal to abstract, impartial and universal principles

> misses out on a good deal of what actually concerns people in their everyday moral lives. [It] ignores the roles played by special relationships, emotional bonds, discussion and compromise in our everyday moral lives. It misses out on an important fact well recognised by most women, and now by many advocates of feminist ethics generally; that a good deal of our moral lives is concerned with the particular personal relationships we share with others. Personal, caring relationships, with all their individuality, subtlety and complexity, are the cornerstones of our moral lives.[14]

In her *In a Different Voice: Psychological Theory and Women's Development*,[15] Gilligan seeks to show that women's moral values tend to emphasis responsibility, whereas men stress rights. In Gilligan's view, women look to context, while men appeal to neutral, abstract notions of justice. She speaks of a female 'ethic of care' deriving this from nurturing.

Applied ethics

As noted earlier, applied or practical ethics is about finding ways to apply such general principles and ideas of ethical decision-making and ethical living to the concrete contexts of particular professional activity: of medicine, business, scientific research, education or law. And this includes practical systems of education, dispute adjudication and regulation aiming to encourage, develop or apply appropriate values in the disciplines in question.

Traditional legal education includes some final-year study in professional ethics. But as suggested earlier, there are good grounds for seeing this as all too often too little and too late. As Nader and Smith note:

> Robert Granfield, a sociologist who teaches at the University of Denver, wrote a book published in 1992 titled *Making Elite Lawyers* about what happens to Harvard Law School students between their acceptance and the time they graduate.
>
> Granfield says that legal education often turns idealists into amoral pragmatists: 'A lot of people who go into law school have a strong sense of right and wrong and a belief in moral truths. Those values are destroyed in law school, where students are taught that there is no right and no wrong and where such idealistic, big-picture concepts get usurped. The way the majority of students deal with this is to become cynical. They actually come to disdain right versus wrong thinking as unprofessional and naive.'
>
> Professor H Richard Uviller teaches legal ethics at Columbia University

> School of Law. He says: 'Professors try to keep the ideals of the students high. But when they learn that their highest obligation is to the client, students begin to think of the rest of it [duty to society and the justice system] as wimpy and soft and not what clients respect. One reason they get that message is because it is true.' And many professors do not attempt to instil idealism in their students. They come from corporate law backgrounds and, indeed, some law professors maintain lucrative corporate law practices on the side. Others, even critics of corporate power, are simply cynics who see little chance of reforming the legal profession. Such cynicism is infectious and convinces many students, often burdened with large tuition debts in the tens of thousands of dollars, to avoid idealistic pursuits and simply maximise income.[16]

To the extent that this sort of situation still prevails, many different things can and should be done to more effectively address it. But one small step is to encourage students' consideration of broad ethical issues, ideas, theories and possibilities from the start of their legal education, moving from general metaethical and normative ethical ideas to issues of applied ethics, rather than merely tacking some applied ethical considerations on at the end.

Much of the rest of this book is about applied ethics; in particular, it concerns the ethics of legal decision-making at all levels, from the formulation and enactment of legislation, to the creation of law by judges and the day-to-day operation of the legal system, including the behaviour of legal professionals.

Discussion topics

1. What is utilitarianism? Include reference to 'act' and 'rule' versions. What are their strengths and weaknesses? How are they related to law?
2. Briefly explain the ethical ideas of Kant, Ross and Aristotle.
3. What is feminist ethics?
4. What is applied ethics? What are lawyers' ethical responsibilities?

Additional resources

S. Mann, *Economics, Business Ethics and Law*, Lawbook Co, Sydney, 2003, ch 3.

R. Nader and W.J. Smith, *No Contest*, Random House, New York, 1996.

Chapter 10

NATURAL LAW AND LEGAL POSITIVISM

This chapter considers some major jurisprudential themes, issues, schools and debates. The first half of the chapter is concerned with natural law and the positivist backlash. In the second half, we focus upon the influential ideas of John Rawls, Robert Nozick and Richard Posner, concerning distributive justice, or the appropriate, fair and just distribution of resources.

Natural law and legal positivism are the most commonly discussed schools of jurisprudence, so it is important to have a basic grasp of their origins, scope and rationales. But it must be emphasised that these are far from being the only two strands of legal theory. Many others exist, including sociological, psychological, Marxist, Critical Realist, postmodernist and feminist theories.

Several other preliminary points must be made. First, as will become clear later in the chapter, when discussing the post–World War II debates between legal positivists and naturalists, the two approaches are not opposites, as they are often portrayed, but can find substantial agreement. Second, it is better to begin with natural law, as the older theory, and to examine the positivist reaction (backlash) to the evolution of naturalism. Third, natural law has taken on many different legal and political hues over the centuries.

In particular, during the great capitalist revolutions (the English, American and French) of the 17th and 18th centuries, natural law produced a revolutionary emphasis on basic rights and human equality. The reaction that this provoked led to the positivist response of Jeremy Bentham and John Austin in the 19th century. The legal

positivists insisted that law had to be defined as the commands of the legitimate or effective government in power, freed of all other considerations, such as morality, social justice, politics and economics.

Natural law

Natural law theorists argue that society's laws are derived from, or must be derived from, a higher source than the promulgations of governments. The definition given by Freeman in *Lloyd's Introduction to Jurisprudence* is a good starting point: 'the essence of natural law may be said to lie in the constant assertion that there are objective moral principles which depend upon the nature of the universe and which can be discovered by reason'.[1]

Natural law can be defined in terms of:

(a) nature;
(b) ethical or moral considerations;
(c) religious precepts;
(d) social contract;
(e) human reason.

These approaches embody various lines of reasoning, including:

- *Proceeding from what 'is' to what 'ought' to be*: that is, deriving principles from observed human nature or society. This runs into the criticism, first articulated by Hume, that it is an unacceptable logical inference to jump from factual observations to normative propositions.
- *Viewing nature and humanity teleologically*: that is, viewing them as tending toward pre-determined ends (for example, acorns grow into oaks). The teleological view sees humanity as having ends that can be ascertained by reflecting on its nature and its needs.
- *Asserting that propositions of natural law are self-evident*: that is, as given by God, as derived from the common features of human societies, or as having been determined by economic and political struggle. Hence the American Declaration of Independence in 1776 proclaimed: 'every man is endowed by his Creator with inalienable rights, and that among these are life, liberty and the pursuit of happiness'. These were described as 'self-evident' but in reality arose out of the struggle of the aspiring American capitalist class

against British colonial domination.

- *Conceiving natural law as having a variable content, according to time, place and circumstance*: This is an attempt to overcome the fact that natural law theory is capable of embracing varied social orders. Thus, in their time, natural law thinkers tended to see both slavery and feudalism as moral and natural.

These strands are not exhaustive and also overlap. But they share some immediate implications. One is that bad laws may be assessed as defective and perhaps not laws at all. Another is that there are limits on the obligation to obey official authority.

This is precisely why Bentham, in his *Anarchical Fallacies* and *A Fragment on Government*, railed against Blackstone's *Commentaries* for suggesting that the common law embodied natural rights. Bentham wrote: '[T]he natural tendency of such doctrine is to impel a man, by the force of conscience, to rise up in arms against any law whatever that he happens not to like.'[2] This was written in 1776, the year of the American Revolution, which claimed a right to overthrow British rule.

Broadly speaking, there are four epochs of natural law:

1 Ancient Graeco-Roman theory;
2 the Judaeo-Christian tradition;
3 the Age of Reason;
4 the Post–World War II revival.

The Ancient Greeks and Romans

The two most important figures were Plato (427–347 BC) and his student Aristotle (384–322 BC). Plato also referred to the thought of Socrates, who had been his teacher.

They wrote and taught in a city-state system that rested on slavery, which created the economic conditions for reflective thought by a privileged layer. They were generally rationalist in their approach. The ultimate source of concepts of good or evil, whether divine or natural, was not important to their legal theory.

Plato's 'ideal' means of conducting society were open to suitably trained human reason. Hence, his 'model' of a 'philosopher king' with little role for a legal code. But he contended that laws should not only compel; they should also persuade and educate in virtue.

On the obligation to obey, there is an apparent contradiction between the *Apology* and the *Crito*. In the *Apology*, Plato drew on the experience of Socrates' trial for impiety, corruption of youth and, in effect, sedition. He asserted that the state has no right to demand that a person commit evil, with the only honourable course being refusal.

However, in the *Crito*, he said an individual has an obligation to obey a law that wrongs him or her, as distinct from requiring wrongs to be committed by the individual. In other words, individuals cannot validly be compelled to wrong others, but must submit to the injustice of the state themselves.

Plato gave three reasons to obey (all of which feature throughout the history of jurisprudence): (1) parent–child analogy; (2) social contract (those who disagree can leave); (3) unrest will destroy the social fabric.

Without dealing with these in detail, it may be said that each is dubious.

1 The parent–child analogy is much overstretched, not least because societies are historically forged and contain conflicting interests, pursued by adults, whereas children are largely dependent upon their parents.
2 The right to leave society is often unrealistic, as it is today for many refugees.
3 The social fabric may be the cause of intolerable injustice and hence discontent.

Aristotle took a teleological approach: human beings have an inherent potential for good, the achievement of which is it is the proper function of the state to facilitate. He regarded humans as 'political animals' combining for mutual life in societies. He postulated the existence of universal justice higher than that expressed in 'good' laws. Aristotle said little on obligation to obey, seemingly agreeing with Plato.

Aristotle tutored Alexander the Great, who undertook military expansion of the ancient Hellenistic world. This was reflected in the rise of Stoic philosophy, which taught that there is a rationally observable higher order, a cosmic reason, which may be appreciated by all peoples, not just a 'civilised few'.

This universalism held attractions for the theorists of the Roman Empire, with its vast and diverse territories. This can be seen in the

work of Cicero (106–43 BC), the most important pre-Christian Roman legal theorist. He defined natural law as 'right reason in agreement with nature'. He also postulated the striking down of positive laws that contravened natural law.

He drew a distinction between positive law, *lex vulgus* – essentially an exercise of political power that might or might not be appropriate – and the divine law, or *lex caelestis*. The latter was accessible to rational insight and inquiry, which would produce natural law, the *lex naturae* – the proper model for making laws.

In Roman practice, this notion of natural law found expression in the concept of *jus gentium*, a body of legal principles common to all peoples of the world, as compared to the *jus civile*, the particular law of a given state. Conquest and trade necessitated the development of law that could be applied to foreigners, hence the assertion of universal law.

The Judaeo-Christian tradition

The final phase of Graeco-Roman jurisprudence was readily adaptable to the adoption of Christianity as the official religion of the Roman Empire by Emperor Constantine in 312 AD. This required a fusion of the Hellenistic and Judaeo-Christian traditions, which continues to influence jurisprudence today.

This tradition is based on an unequivocal and more absolute assertion of a higher, divine authority, with a detailed legal code, as set out, for example, in Exodus 20:1 to 22:17. Like natural law theory as a whole, this can be used to either sanctify an existing order as 'the will of God' or justify disobedience by appeal to a higher authority.

Certainly, Christ was crucified by the Roman state for encouraging revolt against slavery, but after 312 the Church became incorporated into the Roman order, wherein it was used to legitimise the Empire. This also happened with the emergence of the feudal order after the Dark Ages that followed the fall of the Roman Empire.

Christian theory tends to have particular problems with change, since the written scripture is meant to be the eternal law of God, expressed in very concrete prescriptions. As we shall see, Aquinas had difficulties in enunciating a theory to allow for changing economic requirements.

The first early theorist was St Augustine of Hippo (345–430), who himself converted to Christianity and wrote shortly after

Constantine's adoption of Christian doctrine. His best-known assertion is that 'an unjust law is no law'. Citing Cicero, he equated unjust governments with criminal gangs, or pirates – able to have their way only through force. This did not necessarily mean a right to rebel. An unjust law could be coercively enforced, but would not have any moral force. Anything just in positive law (*lex temporalis*) would derive from eternal law (*lex aeterna*).

Augustine had a minimalist view of positive law, limiting it to the coercive discouragement of sin. By the 13th century St Thomas Aquinas had radically altered this latter view. He was writing nearly 900 years later as part of the consolidation of feudalism after the Dark Ages. By then the works of Aristotle had been rediscovered. Like Aristotle, Aquinas considered that positive law plays a proper and 'natural' role in political and social life. In the words of one writer, he afforded the law a greater dignity than suppressing sin. By this stage, the Church was thoroughly institutionalised as part of the state. Arguably, this was a greater influence than the rediscovery of Aristotle.

Aquinas postulated four types of law:

1 *lex aeterna* (eternal Will of God);
2 *lex divine* (divine law, as revealed by scripture);
3 *lex naturalis* (natural law, the fruit of rational human observation);
4 *lex humana* (positive law, 'good' as far as it rests on these foundations).

The idea here is that God's eternal law directs all things in the universe towards their own particular 'goods'. So, as Waluchow notes, to say that X is good for S is to say (a) X is an end towards which S is naturally inclined or (b) X is a means of achieving Y and Y is an end towards which S is naturally inclined.[3]

But whereas in a non-rational thing (like a rock, a plant or a lower animal) such law is manifest as physical necessity (for example, gravity) or as instinct, rational creatures, like humans, can understand such law through reason and choose to act upon it through free will. Natural law is impressed upon us as characteristically human inclinations or tendencies, including the use of reason to understand such tendencies. Primary principles of natural law including the pursuit of self-preservation, the production, care and education of children, pursuit of knowledge and sociability are self-evident to rational inquiry as major driving forces of human life. Secondary principles, including basic tenets of the Roman Catholic religion and acceptance of rigidly hierarchically structured social relations, are supposedly log-

ically entailed by the primary ones, but are not obvious to all. They can be obscured by the social corruption of basic instinct, leading to attachment to such things as homosexuality, sex outside marriage, abortion and revolutionary challenge to a hierarchical social order.[4]

At this point reason needs assistance from both divine law, in the form of moral standards explicitly revealed by God in sacred texts, and human law, restraining those who have not been properly trained in habits of virtue. Such human law quite simply forbids, and punishes, transgressions of natural (moral) law, and aims to encourage and support adherence to such moral law. And any human law that radically departs from natural law is not a true law at all.

According to the Thomist view, tyrannical law made contrary to reason is a perversion of law. In those circumstances, the moral obligation to obey fails, unless greater scandal would result from disobedience. That is, some degree of unjust government should be tolerated for fear of bringing worse things by rebellion or disobedience, but there are limits. Tarquinius, the last king of ancient Rome, was cited as an example of a properly deposed tyrant.

In the battle against the rising capitalist class, this doctrine became more openly reactionary. In the Counter-Reformation, later Thomist thinkers insisted that to challenge positive law was to sin against the eternal law of God.[5] Thus, before they were ousted by Cromwell, the Stuarts claimed 'the Divine Right of Kings'.

The Age of Reason

The Age of Reason, which culminated in the Enlightenment of the 18th century, was bound up with the expansion of man's geographic, cultural and intellectual horizons over the two previous centuries. Some of the factors involved included: great scientific advances; the discovery of the New World; the expansion of world trade; the emergence of nation-states, first in Britain and then the Netherlands, the United States and France.

We cannot here review these developments and their impact in detail. Suffice to say that the unchallengeable authority of the Church had been slowly eroding since Copernicus, followed by Galileo, proved that the earth was not the centre of the universe but revolved around the sun. Giordano Bruno went further and postulated many suns and an infinite universe. For this he was burnt at the stake in Rome on 16 February 1600.

The great advances in science were reflected in philosophy, for example, in Locke's (1632–1704) *Essay Concerning Human Understanding*, which repudiated the concept of innate ideas, given by God, and established the objective source of thought in sensations derived from the external world.

Man's thinking, and therefore his moral character, was, in the final analysis, a reflexive product of the material environment in which he lived. Contained within this conception was a profoundly subversive notion: the nature of humanity could be changed and improved upon by changing and improving the social environment. Humanity was no longer cursed forever by original sin.

How was life to be improved? Through the invincible power of human reason. The motto of the Enlightenment, as Kant (1724–1804) wrote, was 'Dare to know!'

Legal theory likewise sought to locate the state's authority and legitimacy in human, rather than divine, sources. This led to a recasting of the naturalist analysis in revived forms of the social contract, in some ways reminiscent of Plato. This is common to the three principal jurisprudential figures of the 17th and 18th centuries, Hobbes, Locke and Rousseau. However their theories were very different, reflecting different historic and political circumstances.

Thomas Hobbes (1588–1679) wrote his principal work, *Leviathan*, in 1651, in reaction to the Cromwellian revolution in Britain, which had seen a civil war, a king lose his head, the establishment of a parliamentary dictatorship and the emergence of radical egalitarian tendencies such as the Levellers.

He argued that the proper purpose of government and law was primarily to guarantee peace and order. In a time of war, he wrote, 'the life of man is solitary, poor, nasty, brutish and short'.[6] He set out two basic principles:

1 People should strive for peace but may resort to self-defence when the endeavour proves impossible.
2 People should be satisfied with as much liberty as they are willing to allow to others.

For Hobbes, the obligation to obey only ceased when the sovereign fails to maintain the order that is the fundamental term of the social contract. Then the individual right of self-defence will abrogate the duty of obedience owed to the ruler.

Whereas Hobbes was the theorist of the post–Cromwell Restoration, John Locke (1632–1704) provided the theoretical underpinning for the 1689 'Glorious Revolution', in which the parliament and the emergent capitalist class re-established their supremacy by overthrowing James II and replacing him by Protestants, William II and Mary II of Orange.

Locke is a seminal figure in political and legal philosophy, with two principal, interlinked influences. He provided initial rationales for the concepts of liberty and private property, while leaving the underlying tensions between the two unresolved.

His rationale for the Settlement between the monarchy and the parliament had two key planks:

1 rejection of any 'absolute' power in favour of a limited sovereign;
2 all individuals have 'natural rights' to life, health, liberty and property.

If these principles were transgressed, people had a right to resume their original liberty. Thus, he proclaimed a right of revolution, although he was not an enemy of political authority. He postulated the existence of 'tacit and scarce avoidable consent' as well as express consent.[7] Merely by remaining within a state, people tacitly consented to obey its laws because they benefited from the actions of its sovereign. He declared that to disturb government was also to breach the law of nature – it could only be justified when the sovereign had betrayed his trust.

Locke helped forge the necessary ideological weapons for the emergence of a new capitalist society. One of the most important battles in the development of capitalism was the establishment of *exclusive* property rights, above all in land, to prevail over the common property rights that had played such a central role in the lives of the peasantry under feudalism.

This new form of property had to establish itself against the conception that land should be held in common and its fruits available to all. The forms of property, based on exclusion, which are considered as emanating from human nature today, were once regarded as so 'unnatural' that they had to be argued for. Locke wrote:

> But this [that the earth was given to mankind in common] being supposed, it seems to some a very great difficulty, how any one should ever come to have a *Property* in any thing … I shall endeavour to show, how Men might come to have *property* in several parts of that

> which God gave to Mankind in common, and that without any express Compact of the Commoners.[8]

Locke could be characterised as the advocate of propertied revolution. By insisting on the inviolability of property against tyranny, his was the classic bourgeois outlook. This notion found its way into judicial judgments. Take, for example, this remark of Pratt CJ in *Entinck v Carrington*: 'The great end for which men entered into society was to secure their property.'[9]

However, Locke's conception of property was more complex than simply privately held assets. It included the right to the fruits of one's own labour. He deplored the growth of inequality and espoused a right to physical subsistence, even where it cut across property rights. Locke argued that if a man insisted on the market price for food for a man dying of hunger, he was guilty of murder.[10]

In many respects, the social contract approach to natural law culminated in Jean-Jacques Rousseau (1712–78), who wrote in the lead-up to the French Revolution. Unlike Locke, he postulated rights to life and liberty, but not property. In fact, private property was at the centre of his famous observation, uttered at the outset of *The Social Contract*, that 'human beings are born free but are everywhere in chains'.[11] The revolutionary implications of this insight found expression in his *Discourse on the Origin and Foundation of Inequality among Men*, published in 1755. Property, he explained, was not a natural attribute of human existence. In his natural state, man did not have property. It is the product of the growth of civilisation, which, once having come into existence, destroys man's humanity and enslaves him.

> The first man who, having fenced off a plot of land, thought of saying, 'This is mine', and found people simple enough to believe him was the real founder of civil society. How many crimes, wars, murders, how many miseries and horrors might the human race have been spared by the one who, upon pulling up the stakes or filling in the ditch, had shouted to his fellow men, 'Beware of listening to this impostor; you are lost, if you forget that the fruits of the earth belong to all and that the earth belongs to no-one.'[12]

As there was once no property, so was there once no inequality. Like property out of which it develops, inequality is a product of civilisation. The poor are oppressed by the power of property. Those who possess property are morally and intellectually disfigured by the struggle to obtain, keep and augment it.

The background to the positivist reaction

Writing before the French Revolution, David Hume (1711–76) was among the Enlightenment figures whose rationalism and secularism came into conflict with the idea of a universal natural law common to all mankind.

He declared it was a leap of logic to deduce 'ought' from 'is', that is, to derive the normative prescriptions of natural law from descriptive observations of humanity and society as it is.[13] In addition, Hume asserted that governments and laws were the creation of men and reflected 'human interests' and 'conventions', not the laws of nature.

Nonetheless, Hume developed an empirical view of the principles of justice, describing them as 'natural laws' because 'they are as old and universal as society and the human species, but prior to government and positive law'. He said observation of the rules of justice was 'palpable and evident, even to the most rude and uncultivated of the human race'. So perhaps the separation of man and nature is not so simple?[14]

In response to the American and French Revolutions, western jurisprudence took a sharp turn, particularly in England. This took two forms, one personified by Burke and Blackstone and the other by Bentham and Austin.

Edmund Burke appealed to natural law to attack French egalitarianism and to support the wage system. He was a conservative whose heritage is often claimed today by Tory parties.

He appealed to 'the eternal principles of truth and justice' to denounce arbitrary rule and interference with the rights of property and employers. He railed against the Speenhamland system of paying labourers a supplement related to the cost of bread and the size of their families. This was because the wage relation was part of a natural 'chain of subordination'. He equated the capitalist market with the divine and natural order and asserted that capitalism 'had in fact been the traditional order in England for a whole century'.[15]

Similarly, in his *Commentaries*, Blackstone appealed to natural law to sanctify the English common law. For this, as mentioned earlier, Bentham denounced Blackstone because the natural law doctrine tends to impel people, by force of conscience, to rise up against laws they do not like. Natural law was 'nonsense on stilts'.[16]

Classical positivism

The founder of classical legal positivism is Jeremy Bentham (1748–1832), whose ideas were developed, arguably to their detriment, by John Austin (1790–1859). Positivist theories describe law as it is in a given time and place, by reference to formal, rather than to moral or ethical, criteria of identification.

This has been the dominant school since the early 19th century. It also accords with the demand for a 'practical' definition of law. But as we shall see, that quest is fraught with difficulties.

Bentham was far from indifferent to the quality of law. He actively sought to correct laws that he regarded as offending against principles of utility. However, he made a distinction between expositorial jurisprudence (what the law is) and censorial jurisprudence (what it ought to be). For the latter, he sought to develop a 'science of legislation'.

Classical positivism had two main features: a separation of law from morality and other factors, such as economics, and an attempt to define law by a command theory.

The command theory as suggested by Bentham and developed by Austin was simplistic and divorced from the underlying forces giving rise to a legal system. It postulated three requirements: a command, a sovereign, and a sanction.

Bentham defined a law as follows: 'an assemblage of signs declarative of a volition conceived or adopted by the sovereign in a state, concerning the conduct to be observed by persons, who are or are supposed to be subject to his power: such volition trusting for its accomplishment to the expectation of certain events'.[17]

But not all laws are simply orders, even on a positivist view. For example, laws facilitate the making of contracts and wills. Another obvious difficulty was that by the early 19th century, the sovereign power, at least in England, was limited and somewhat dispersed. There was a separation of powers, and also local government.

Bentham was anxious to avoid the assertion of a right to rule; hence he focused on the so-called fact of rulership. He defined a sovereign as 'any person or assemblage of persons to whose will a whole political community are (*no matter on what account*) supposed to be in a disposition to pay obedience'. (our italics)

Thus, the basis of authority was an actual or supposed habit of obedience, no matter why. This view not only leaves room for tyrants

and dictators, it cannot account for revolutions, military coups, acts of secession, etc, which erupt when people are no longer disposed to be obedient.

Bentham depicted sanctions in terms of pain and pleasure. Fear of pain was a coercive sanction; expectation of pleasure was an alluring sanction. He admitted other motivations for compliance with law, including physical, political, moral and religious sanctions, but argued that a sanction imposed by a sovereign was a definitive characteristic of law.

Among other things, the command theory cannot explain legitimacy. Can obedience be reduced to habit? Do not reason and agreement play a role? Is mere coercion sufficient to retain stability? In the words of McCoubrey and White: 'Laws are much more complex than simple orders, and a sovereign body is not just a glorified sergeant-major.'[18]

Modern positivism and the revival of natural law

Now we turn to the 20th-century debate between modified positivism and revived natural law, notably the so-called Hart–Fuller debate, augmented by Finnis, a contemporary naturalist. These theorists emerged in the context of the post–World War II restabilisation after the victory of the US and Britain over Germany and the beginning of the Cold War. Their discussion on the nature of law was also motivated by the experiences under the Nazis.

The perceived abuses of legal process inherent in the Nazi and Stalinist regimes led to a questioning of the validity of formalist legal theories. Positivism's exclusion of questions of morals, ethics, human rights and political considerations from the realm of jurisprudence – even if they were to be examined in other spheres or disciplines – seemed to ignore matters that should be central to the nature of law in the 20th century.

On one side of the debate, HLA Hart sought to develop a more sophisticated form of positivism, avoiding the rigidity and narrowness of the Austin–Bentham command theory by outlining a less precise notion of the concept of law.[19] One of Hart's main contributions is said to be his enunciation of a scheme of primary and secondary rules. He suggested three categories of secondary rules: rules of recognition, rules of change and rules of adjudication.

Hart's rules of recognition were ultimate rules underpinning a legal system. However, he did not provide a clear definition of the rule or rules of recognition, instead suggesting that common acceptance by officials was sufficient.[20] He illustrated his conception by the simplistic, mythical example of a tyrant, Rex I – in effect, a crowned gunman, whose son Rex II succeeds him and is accepted as legitimate on the basis of the rule of succession by the eldest son. In Hart's view, this was adequate to constitute a formal right to rule. This seems a poor example in the 20th century, divorced as it is from any conception of popular support or acceptance.

Later, Hart enunciated what he called a 'minimum content of natural law'. Despite his formalistic legal positivism, he asserted that there was a minimal, or perhaps, an essential moral content to law. He claimed there were 'truisms concerning human nature' that do form a common element in the law of all societies. His five 'truisms' – approximate equality, human vulnerability, limited altruism, limited resources, limited understanding and strength of will – are highly debatable. For example, what is 'approximate equality'? It is certainly not genuine social equality, nor a right of any kind. Perhaps, it is based on a rough political assessment of what will be tolerated without social unrest.

The Hart versus Fuller debate took place on the question of the validity of some or all Nazi laws and judicial decisions. Hart's argument was broadly that the Nazi laws, however oppressive or immoral, were in accord with the rule of recognition and must be considered to be law.

If Nazi informers were to be punished, it would be preferable to openly state that as a matter of policy and enact retrospective legislation to do so, rather than infringe on the distinction between law and morality. Thus, to preserve his positivist conception, Hart was prepared to embrace backdated criminal laws, despite a centuries-old principle abhorring such retroactive measures as permitting arbitrary rule.

Lon Fuller was professor of general jurisprudence at Harvard from 1948 to 1972. He sought to answer the issues raised by the belatedly recognised totalitarianism of the 1930s by advancing a theory that he labelled 'procedural naturalism' in an effort to define the minimal requirements for a recognisable legal system.

There is debate whether the Nazi regime was a 'rule of law' and some commentators have observed that Hitler's administration would have passed nearly all the requirements specified by Fuller. He insisted that his was a natural law theory but, arguably, Fuller's scheme is not correctly termed 'naturalist' in that it is primarily concerned with the

minimal procedural prerequisites for a legal system, not the substantive content of a legal system.

In *Morality of Law* he explains the narrower task that he assigned himself, although he saw it as a larger issue: 'There is little recognition … of a much larger problem, that of clarifying the directions of human effort essential to maintain any system of law, even one whose ultimate objectives may be regarded as mistaken or evil.'[21]

Fuller rejected the Christian doctrines of natural law and sought to apply human reason to discover 'principles of social order which will enable men to attain a satisfactory life in common'. His 'morality of duty' sought to lay down eight 'basic rules without which ordered society is impossible, or without which an ordered society directed toward certain specific goals must fail of its mark.'[22] Laws had to be (1) general; (2) publicly promulgated; (3) sufficiently prospective; (4) clear and intelligible; (5) free of contradictions; (6) sufficiently constant; (7) possible to obey; (8) administered congruently with their wording.

This raises many questions. What is an 'ordered society'? Whose order? Are the procedural means, even in this minimal sense, so easily separated from the ends of society? Fuller himself argued that the procedural requirements will 'affect and limit the substantive aims that can be achieved through law'. He wrote of their 'reciprocal influence'.

Fuller asserted one 'imperious' tenet of substantive naturalism in the maintenance of 'channels of communication' between people and peoples.[23] This attempt to define, perhaps, a minimalist content of natural law sits as incongruously with Fuller's proceduralism as does Hart's rather different 'minimum content' with Hart's separation of law and morality.

There is no doubt about John Finnis' credentials as a natural law scholar. He set the classical natural law concerns of Aristotle and St Thomas Aquinas in the contemporary language of 'natural rights'.

He took issue with David Hume's 'is' and 'ought' deductive fallacy by denying that natural lawyers have ever sought to derive ethical norms from facts, that is, from simple observation of human conduct. He argued that people have an 'internal' perspective of their aspirations and nature and from this it is possible to extrapolate an understanding of the 'good life' for humanity in general.

He claimed to derive 'seven basic forms of human good' from a survey of anthropological investigations: life, knowledge, play, aesthetic experience, sociability or friendship, practical reasonableness, and religion.[24] Obvious questions arise. Are these exhaustive?

Essentially, Finnis argued that they are. He said there are countless objectives and forms of good but on analysis they are always found to be one or a combination of the seven basic forms of good.

The list has some troublesome omissions. What about political freedom? Equality? Freedom from exploitation? Indeed, the list seems culturally specific to contemporary free market capitalism.

For example, what is 'practical reasonableness'? Finnis wrote of 'a measure of effective freedom' and 'genuine realisation of one's own freely ordered evaluations, preferences, hopes and self-determination'. This is quite relative and individualist. There are none of the inalienable rights of liberty, fraternity and equality invoked by the leaders of the French Revolution. There is no right to 'happiness', let alone insurrection proclaimed by the American Revolution.

And how can such conceptions, even as vaguely phrased, be realised outside economic welfare and freedom from material want? There is no mention of this under the heading of life, only references to famine relief and keeping children alive until they can fend for themselves. This is not only an extremely minimalist notion, it is tied to family responsibility, not social responsibility.

Notably, Finnis defended private property. He argued that 'the good of personal autonomy in community' suggests private property as a requirement of justice. A socialist, following the observation first articulated by Rousseau, would suggest the opposite: that private ownership of the means of production, resting on wage slavery, is incompatible with genuine personal autonomy.

Finnis, like Aquinas, greatly qualified the right to defy an unjust law. In effect, he argued that in disobeying a law, even a bad law, a person places at risk the whole legal system and that therefore there may be a 'collateral' moral obligation to obey such a law.[25]

Notwithstanding these and other problems, which are arguably inherent in contemporary natural law theory, conceptions of natural law have been broadly influential since the middle of the 20th century, notably in international law. This can be seen in the development of international conventions and treaties, such as the Geneva Conventions on the conduct of war (1949), the Universal Declaration of Human Rights (1948), the Refugee Convention (1951), the International Covenant on Civil and Political Rights (1966), the International Convention on the Elimination of All Forms of Racial Discrimination (1966) and the International Covenant on Economic, Social and Cultural Rights (1966).

Natural law and science

It is important to see that while natural law theory underwent significant development in a feudal and Catholic Christian context, it is quite possible to develop secular natural law theories based upon objective scientific truths about what is good for all human beings. In particular, a number of contemporary theorists have followed Plato in focusing upon medical considerations of physical and mental health as crucially bridging the gap between fact and value and thus providing a solid foundation for a modernised, secular natural law theory.

The crucial point about natural law theory is the idea that rational faculties of logic, observation, experiment and calculation can allow us to identify – and perhaps quantify – objective facts about what is good for human beings. And given that biological, psychological and social science show that all humans have a range of basic needs that have to be satisfied in order for them to remain healthy and fully realise their human potentials of mental and physical growth and development, this would seem to be a logical starting point for contemporary natural law theory. That which promotes such healthy development is good, that which obstructs it, bad.

Here is a possible rational foundation for particular (legally supported) rights and duties. With basic needs recognised as rights of all, crucial social and legal questions revolve upon effective means of achieving or realising such rights, equally and for all.

Social contract theory

We have already briefly touched upon the origin of social contract theories in the 17th century. The central idea here is the derivation of moral principles (of necessary respect for the interests of others) from rational self-interest. And this is still the starting point of contemporary social contract theories of morality as developed by John Rawls, David Gauthier and others.

As Waluchow notes, on this sort of account, morality is simply a set of standards that would be agreed by rationally self-interested individuals clever enough to see the benefits of living in a society where such standards are observed. So the standards are essentially derived from prudence or self-interest and can vary considerably depending upon the priorities of the individuals concerned.[26]

The Prisoner's Dilemma has figured prominently in recent discussions of such social contract morality. The idea here is that of two

prisoners separately interrogated about a crime. If one confesses and the other doesn't the confessor gets a light sentence and the non-confessor a very heavy one. If both confess, each gets a sentence of intermediate length, if both keep quiet they each get a sentence only slightly longer than that of the single confessor.

		Their Choice	
		Confess	Dont Confess
Your Choice	Confess	5, 5	1, 10
	Dont Confess	10, 1	2, 2

(The first number gives your gaol time, the second one their gaol time, given the choices in question.)[27]

Ideally, both should say nothing. But since neither can be sure of what the other will do, both will reason as follows: If he confesses, I better do so. And if he doesn't confess, I still better do so. So each ends up doing five years instead of two. If each knew the other could be trusted to co-operate with them, then each could keep quiet. This is supposedly analogous to a state of nature situation. Morality is about avoiding '5,5'-type situations by 'contracting together' to create situations in which it is rational to expect co-operation.

Such situations involve the creation of alliances of co-operators, which effectively exclude non-co-operators from the benefits of co-operation. Eventually all, or most, come to see the benefits of co-operation. Waluchow offers the beginning of an analysis of the basic rules of such co-operation based upon the Minimax Relative Concession Principle.[28] The idea is that acceptance of this or similar rules defines a particular 'moral community'. People share the benefits of collaboration through following such rules.

We can see the ideological function of such a theory in sustaining contemporary social power structures. For it could be interpreted to suggest that we have all freely contracted to current social and legal arrangements and do indeed all benefit equally, or at least fairly, through following such established rules, that the only alternative to such arrangements is a chaotic 'state of nature'. In fact, current arrangements are far from fair and mutually beneficial, far from being the product of meaningful negotiation and agreement between all parties, and far from being the only social arrangements possible.

John Rawls and distributive justice

We conclude this chapter with brief consideration of some major contemporary approaches and debates about the appropriate, fair or just distribution of valuable resources. In particular we focus upon the influential contributions of Rawls, Nozick and Posner.

John Rawls stands in the liberal tradition. He arguably represents a high-water mark in 20th century, postwar liberalism. His two primary works, *A Theory of Justice* (1972) and *Political Liberalism* (1993), bear the imprint of the radicalisation and working-class upsurges of the late 1960s and early 1970s. He enunciated two principles of justice: liberty, and reasonable social and economic equality. The meaning of these propositions can be summarised as follows:

1 Maximisation of liberty, subject only to such constraints as necessary to protect liberty itself.
2 Equality in the basic liberties of social life and also the distribution of other social goods, subject only to a 'difference principle' or 'compensating benefits'. Rawls argued that inequalities are permissible if they produce the greatest benefit for those least well off.
3 'Fair equality of opportunity', which requires the elimination of privileges based on birth or wealth.[29]

The 'difference principle' not only assumes the continued existence of inequality but also justifies it. It goes a long way to embracing a utilitarian justification of inequality on the grounds of benefit to the majority. Unlike utilitarians, Rawls did hold some rights to be reasonably absolute – those concerned with liberty. But, for him, equality was a lesser priority 'lexically' (a matter of language only?) than liberty, and merely a greater lexical priority than efficiency and the welfare of the majority.

Rawls claimed to 'refurbish' social contract theory. He incorporated a notion of co-operation: the intuitive idea that all benefit from a scheme of co-operation; yet his scheme retained an individualist view of society. He based himself on a rather fantastic notion of 'original position' reached behind a 'veil of ignorance' in which all the participants are stripped of their knowledge of their own personal and social characteristics.[30]

He did not assert that his view was based on reason or nature or empirical proof, but on mutually acceptable ground rules for society.

Rawls claimed that rational individuals in his original position would choose a 'maximin' strategy, aimed at guaranteeing that the worst condition one might find oneself in is the least undesirable of the alternatives. In his *Political Liberalism*, Rawls added three further caveats: reasonable pluralism, the 'fact of oppression', and a 'thin theory' of distribution.[31]

The first demonstrates the minimalist character of his schema: it is based on an 'overlapping consensus' on justice, but beyond that all can disagree. The second seems to justify the oppressive use of state power to maintain shared religious, philosophical and moral understandings in society. The third reveals that only minimal propositions related to wealth and self-respect, upon which unanimous accord is possible, are binding.

Most criticism of Rawls centres on the unreality of his 'original position'. As with social contract theory generally, his view is ahistorical because it ignores the underlying economic, political and social processes by which societies have actually been forged in different historical epochs. More than that, Rawls' conception denies the material (class, economic and social) interests and conflicts that shape the socio-economic structure, as well as the ethnic, religious, gender and ideological differences that are generated by these conflicts. In the final analysis, Rawls' schema is a concentrated model of liberal myths. It reduces members of society to atomised, notionally equal, individuals.

In addition, Rawls seeks to divorce liberty from equality. In reality, liberty (for most people) is illusory without economic equality. Rawls seems to assume a well-off society. He argues that at a 'certain level of wellbeing' citizens would value liberty higher than economic justice. But how is that happy state of prosperity to be achieved for all? Rawls seems to have no answer. As a matter of fact, Rawls' theory appears to highlight this contradiction by introducing a conception of 'worth of liberty' proportional to a person's capacity to advance their claims to liberty. A person's capacity to assert their liberty is likely to depend upon their economic power.

Despite the limitations of Rawls' conception, his views have been condemned by conservative and pro-market theorists. In reaction against the radicalisation of the 1960s and 1970s and the reassertion of basic rights, including distributive rights (reflected in Rawls), two notable countervailing trends emerged, one associated with Robert Nozick and the other with Richard Posner (the economic analysis of law).

Libertarianism: Nozick and the minimal state

Libertarianism and economic liberalism represent minimalist positions in relation to distributive justice. Citing Locke, they both postulate an original right to freedom and property, thus arguing against redistribution and social rights and for the free market. They assert an opposition between equality and freedom: the individual (natural) right to freedom can be limited only for the sake of foreign and domestic peace. For this reason, libertarians consider maintaining public order the state's only legitimate duty.

Robert Nozick's view, enunciated in his 1974 volume *Anarchy, State and Utopia*, is based, to use his words, on the 'classical liberal theory' of a nightwatchman notion of the state (see chapter 11). His view is, in a sense, the sharp revival or reassertion of 19th-century laissez-faire capitalism. His minimal state essentially consists of law and order and protection of private ownership of production. To guard against anarchism, he opposed what he termed an ultra-minimalist state, insisting that the state have a 'monopoly over force' and, in return, provide 'universal protection of rights'. Nozick did so even though this concession offends against his objection to any redistributive function being performed by the state. He challenged the concept of distribution to the extent of rejecting taxation as akin to forced labour. His only 'moral' objection was that compulsory security is necessary to protect the rights and property of all.[32]

Nozick's 'entitlement theory' postulates an inherent right to private and individual ownership of economic goods. He insisted upon the absolutely inviolable character of property rights. But where do these rights come from and how do they arise? Nozick did not argue the legitimacy of the initial appropriation of property, only asserted it as a fact that should be maintained. He wanted society to 'Hold onto the notions of earning, producing, entitlement, desert and so forth.'[33] Yet even Locke, whom Nozick invoked as authority, said the earth was common property (see chapter 8).

Nozick said the right to property is an expression of the right to liberty, yet defined liberty by reference to the right to property. Beside this argument being circular, the more property held by a few, the greater the impact on the liberty of others.

Posner's economic analysis of law

Richard Posner asserted both normative and descriptive propositions: that law *ought* to be concerned with wealth maximisation – which it equates with economic efficiency – and that law as practised *is* actually determined by those values. Posner argued that in order to promote efficiency, courts should mimic the market.[34]

Yet, the market is determined by those who control society's wealth. It allocates on the basis of private profit and power over resources, to the detriment of others. The wealthy few can make their choices at the expense of those who lack the assets to do so. In fact, those who monopolise the wealthy few use their assets to coerce the majority, primarily by way of wage labour.

Posner presents terms such as 'efficiency' and 'optimality' as value-free but the reality is different. Economic decisions portrayed as depending upon willingness to pay in fact depend on capacity to pay. Moreover, the so-called economic conception of man assumes universal self-interest, even greed. It rules out altruism and concern for the state of society and the environment. It can be shown that the entire argument is circular and purely definitional. That is, 'what people want' is defined as what they are assumed to do under the economic model.

A related point can be made. Posner's theory equates individual wealth maximisation with social wealth maximisation, as if the latter were simply an aggregation of the former. As a matter of fact, individual wealth accumulation may be at the expense of social wealth. One can think of the patenting of medical and scientific discoveries. Microsoft and Enron, for example, have arguably dominated technology and markets at considerable social cost.

Posner contended that wealth maximisation serves to 'yoke selfish desires – which in most people are their strongest desires – to the service of other people, and to do so without coercion'.[35] On this basis, Posner claimed that his theory is a moral one. But this 'trickle down' notion of rising affluence lifting all is confounded by the modern reality of ever greater social inequality and social polarisation.

In responding to Posner, Richard Dworkin made a telling point by taking one of Posner's examples and interposing a tyrant to displace the market in allocating resources. The underlying truth in Dworkin's argument is that stripped of all the appealing outward trappings of the supposedly voluntarily entered marketplace, Posner's approach is no different to a tyrant imposing market solutions.

Nevertheless, Posner claimed that all law, even criminal law, could be analysed using his method. Hence his conclusion that 'the prevention of rape is essential to protect the marriage market'.[36] Many of the issues raised by Posner's analysis are explored in detail in chapters 13 to 16.

Discussion topics

- What is natural law theory? What are its strengths and weaknesses?
- What is legal positivism? What are its strengths and weaknesses?
- Briefly summarise and assess the Rawls–Nozick debate about justice.

Additional resources

M. Davies, *Asking the Law Question*, Lawbook Co, Sydney, 1994, Ch. 3.
M. Freeman, *Lloyd's Introduction to Jurisprudence*, 6th edition, Sweet & Maxwell, London, 1994, chs 3, 4 and 6.
H. McCoubrey and N. White, *Textbook on Jurisprudence*, 3rd edition, Blackstone Press, London, 1999, chs 2–5.
R. Wacks, *Swot Jurisprudence*, 5th edition, Blackstone Press, London, 1999, chs 3–5.

Chapter 11

LIBERALISM

This chapter provides an outline of key ideas of neo-liberalism, as the currently dominant form of right-wing liberal political theory, and neo-liberal ideas of the social role of law. This includes the neo-liberal critique of competing left-liberal and socialist political ideas.

Neo-liberalism

Neo-liberals are distinguished by claiming to be more interested in individual rights and freedoms, and in productive efficiency, than in equality. They make much of the apparent incompatibility of liberty and efficiency on the one side and equality on the other, coming down squarely on the side of liberty and efficiency. They argue that because humans have no intrinsic goals or purposes (determined by God), they must be free to choose their own, and act upon such choices with minimum external interference. And the more efficient system of production offers the greater opportunity for free individual choices.

Freedom is all-important to the extent that particular rights and responsibilities arise only out of freely chosen or freely negotiated agreements or contracts between individuals. Individuals have no responsibilities simply by virtue of their social power or position. They incur responsibilities only through such free negotiation, and others can expect to have to pay for the privilege of their care or consideration, as clients of one sort or another.

At the same time, rights to life, to ownership of private property and to take steps to protect one's life and one's private property have

a special significance for neo-liberals. Following a long tradition of right liberalism, they see the state as essentially a construction of free individuals, contracting together to create an effective means of protection of their lives and their property, through the exercise of coercive central power.

Neo-liberals do, typically, subscribe to particular sorts of equality of reward or desert and equality of opportunity. But, as we will see, their interpretations of these ideas are limited and circumscribed and come into contradiction with other aspects of their world view and political program.

A key concept of neo-liberal ethical, political and economic thinking is that of a 'free market' presented as both a fair and just system of distribution of resources – with everyone getting their fair share of the total social product, in proportion to their productive contribution – and an efficient mechanism of organisation of production and distribution, effectively matching supply with demand and maximising the size of the available social product (and thus also of – potential – individual shares of such product) in ways that no other system can achieve.

The marketplace is the arena for the free negotiation of rights and responsibilities, where individuals contract to supply particular goods or services in exchange for what they deem to be appropriate payment or recompense.

The basic neoclassical economic model, favoured by neo-liberals, presupposes 'perfect competition', where many sellers provide identical products, so that none can affect the price of its own product, and many buyers so that none can affect the price they pay; all buyers and sellers have perfect information about all products and prices. It is assumed that if all factors of production – land, labour and capital – receive their appropriate price, a 'general long-run equilibrium', where all available resources are fully and efficiently utilised, will be achieved. Such long-run equilibrium means that the size of each industry has adjusted to the level of demand so that prices correspond to the 'marginal costs' of inputs (land, labour and capital), no business derives 'surplus' profits in excess of the costs of its capital, no productive capacity is wasted and no consumer prepared to pay a price for a good that covers its production costs remains unsatisfied.

Neo-liberals generally see contemporary capitalist society as, by and large, corresponding to such a fair and efficient market model, though the United States is generally seen as much closer to the ideal

than the European Community. And to the extent that the reality falls short of the ideal, the model is taken to provide clear policy recommendations for addressing the shortfalls.

The policy recommendations in question are well known as the program of contemporary neo-liberalism or economic rationalism. They include an emphasis upon the benefits of competition in all areas and at all levels – individuals, businesses, regions, nations – as the path to efficiency, growth and progress. This is taken to support the break-up and privatisation of state monopolies (in the interests of reduced costs to consumers), and winding back of the welfare state and public regulation of corporations in favour of user-pays for services and self-regulation.

The policy recommendations include the prioritisation of low inflation through manipulation of interest rates. In particular, the aim is to stifle a positive feedback of wage increases that drive price increases and in turn further wage increases through increasing interest rates at the first sign of inflationary price rises. At the same time, there is a general aim to keep interest rates low to encourage growth and employment.

So do they include reduced company and personal wealth taxation in favour of broadly based consumption taxes and microeconomic or labour market reforms, moving away from collective bargaining and into individual contracts. This too is supposed to encourage investment and economic growth.

The policy recommendations also include, generally, removal of obstacles to international free trade and investment. Supposedly, in a situation of 'free' world trade, all can benefit from increased productivity and exchange based upon comparative advantage. Similarly, free capital mobility allows machinery and know-how to move where they can most effectively increase productivity, thereby increasing global efficiency.

In neo-liberal theory, such a free market model is underpinned and complemented by the ideas of 'equality of (job) opportunity', representative democracy and a minimal, or nightwatchman, state.

Equality of opportunity refers in this context to removal of legal or customary obstacles to any individuals entering into processes of job selection, and ensuring that there is no discrimination within such selection procedures on grounds that are irrelevant to the appropriate fulfilment of the requirements of the positions in question.

On the one hand, there is a strong right-liberal tradition, contin-

uing in neo-liberalism, of identifying all human beings – unrestricted by central power – as creative, dynamic, innovatory and productive. On the other hand, an equally strong current of thought identifies all humans as, by nature, aggressively competitive, selfish, lazy, greedy creatures, with infinite wants for consumption goods.

At the same time, there are also longstanding right-liberal ideas of humans as strongly 'naturally' motivated to trade and accumulate wealth, and as power-seeking and domineering – at least where they are able to be.

These ideas are – perhaps – reconciled by reference to free markets – as mechanisms uniquely suited for harnessing selfishness, competitiveness, power-seeking and greed, and turning them into the driving forces of creativity, innovation and effort, serving the good of all.

These ideas of 'human nature' are closely tied in with neo-liberal ideas of inequalities of wealth and power as necessary incentives to productive effort, initiative, innovation and best use of available resources to maximise efficiency and output; and ideas of the role of inequalities in fuelling 'aspirations' for more consumption goods, that, in turn, fuel economic growth.

At one end of the scale, substantial rewards are apparently necessary to ensure that the most talented people use their talents to the best advantages of all – or that someone is willing to make the effort and take on the massively onerous responsibilities involved in positions of power and responsibility. At the other, the miseries of poverty and privation are necessary to overcome the inherent sloth, laziness and inertia of some or all human beings.

Profit maximisation by capitalist entrepreneurs drives the operation of market forces that effectively match output to demand and drive ongoing productivity gains through innovation – supposedly achieving output levels, cost reductions and efficient distribution not possible through any other means.

A business that achieves a significant gain in productivity through technical or organisational innovation can sell its goods more cheaply than its competitors, increasing its market share and still make more profit per unit sold, at least until others gain access to the innovation in question. This gives the business super-profits with which to reward the innovators. And this, in turn, encourages further such innovation to 'keep ahead of the game'. Meanwhile, consumers benefit through reduced prices as the innovation is generalised, with competition bringing prices back into line with real production costs.

Representative democracy refers to universal suffrage, supporting periodic elections of representatives to legislative, executive and sometimes also judicial institutions. Thus, (qualified) political experts representing the citizenry actually decide issues and enact laws in the interests of such citizens. But they remain answerable to such citizens at election time, when all such citizens' votes count equally.

All citizens are (equally) free to stand for political office. And equal rights of free speech allow for dissemination and debating of the different political programs of competing candidates and parties. However, political participation – except during elections – is not encouraged. 'Stability and the equilibrium of the system are held as higher values than participation and popular empowerment.'[1]

In the interests of such stability, and efficiency in the exercise of the political managerial skills of the relevant 'experts', neo-liberals tend to prefer the Westminster 'first-past-the-post' system in which the first candidate to achieve enough of the vote not be overtaken by any other, wins the election and becomes the (single) representative for the electorate in question. As Swift points out, such a system tends to revolve around a couple of well-funded parties with similar ideologies, with other perspectives pushed into the sidelines as 'extreme' or 'irrelevant'.[2]

The 'minimal' or 'nightwatchman' state refers to a political system that 'provides sanctions against criminal behaviour (where crime is conceived largely as crime against property), enforces contracts and generally does all and only things that make capitalist markets function well … like establishing a uniform monetary system and controlling the supply of money'.[3] The state should increase the money supply in proportion to the 'natural rate' of growth of the economy (rather than seeking to 'drive' such growth). Central banks should be independent of government control, mandated only to control inflation through interest rate manipulations.

According to neo-liberal theory, these institutions – of the market, democracy and the minimal state – leave all individuals (completely) 'free' to choose what sort of contribution they make to productive (and political) activity. And within such a system, appropriate, fair and just rewards flow from such free decisions, in direct proportion to the effort and sacrifice of those concerned.[4] Such rewards include ownership of private property, both in means of production and consumption, with ownership here understood to include wide-ranging rights to dispose of and benefit from such property (including the right to gift or bequeath it to whomever they choose), endorsed and sanctioned in law.

Individuals should be free to do what they like with their property up until the point where their action interferes with the right of others to similarly 'enjoy' their own private property. At this point, legal action becomes possible to protect the rights of those others. All citizens are equally free to choose to respect the law, and those that fail to do so are equally liable to penalties in proportion to the severity of their crimes.

Again, defenders of the current system judge it to correspond, at least in the developed western states, by and large, to the (political) ideal, with policy recommendations for addressing the divergences.

Here, we consider only one obvious problem with such ideas. The logical development of market success through economies of scale is monopoly and oligopoly, with companies reaping super-profits at others' expense, through excluding others from effective competition. Structural barriers due to high start-up costs, as well as product differentiation through advertising, can function to support oligopoly pricing to the detriment of consumers (and other non-monopoly producers). In theory, neo-liberals are as hostile to private monopolies and oligopolies as to state-controlled ones, and call for elected governments to take action to maintain or restore free competition.

The problem is that such a call for strong action by governments against vast and powerful corporations appears incompatible with the idea of a minimal, nightwatchman state. In face of such incompatibility, neo-liberals tend to fall back upon the idea that market forces still do have the power to overcome such monopoly and oligopoly, so long as governments do not actively support such 'distortions'. Or else, they focus upon the benefits of private oligopoly, in terms of scale economies and global market domination, bringing substantial profits into their home nations, potentially boosting local investment.

In face of the reality of monopoly and oligopoly power of big corporations in the world market, and the unwillingness or inability of governments to challenge such power, it seems that the neo-liberal support for allegedly free existing markets is really support for the freedom of such corporations to continue to expand and accumulate profits without restriction. Indeed, neo-liberal-inspired moves towards deregulation of capital markets around the world seem geared towards removing any remaining obstacles to such processes of concentration and centralisation. This is the central claim of many contemporary critics of neo-liberalism.

The liberal critique of socialism

Socialist and communist ideas can appropriately be characterised as the antithesis of right-liberal ideas, both because they arose originally out of such ideas and because they developed as a radical critique, rejecting key elements of right-liberal theory. Whereas right liberals, including neo-liberals, see inequality of outcome as a natural, acceptable – and desirable – consequence of fair equality of reward, socialists rather see it as a fundamental ethical and social problem.

Socialists see private ownership of productive assets and the operation of market forces – rather than equality of reward – as the primary cause of unacceptable inequalities of wealth, income and power in capitalist society. And they aim to remove such radical inequalities by replacing private with public – collective – ownership of means of production, and planned distribution of goods. In a socialist society such goods would be distributed according to need, aiming to achieve a comprehensive equality of outcome.

Some look to the creation of an initial socialist society, where comprehensive public welfare facilities and state control of productive resources compliment a situation of reward in direct proportion to labour time contributed in the development and application of such resources. Markets are replaced with systems of democratic, participatory planning, with democratic workers' and consumers' councils organising production and distribution at the local level. All jobs are 'balanced for empowerment and desirability', with no-one allowed to do all the 'good' jobs or forced to do all the 'bad' ones. Those who do more of the unpleasant work are rewarded accordingly, with higher wages but never through private ownership of productive resources or increased social power or privilege.

As will be seen in the next chapter, Marxists see such a society as an intermediate step on the path to a communist society – of abundance – in which all receive what they need and contribute according to their abilities. In such a communist society, there is no longer any requirement for laws to protect private property or to enforce punitive sanctions in relation to crimes arising out of the existence – and unequal distribution – of private property.

Right-liberal and social-liberal ideas of democracy involve only periodic or indirect involvement of citizens in the selection of political representatives. Citizens have no democratic control over basic economic decision-making, even about issues that affect them directly,

such as their employment, wages and working conditions. The economic decision-making powers of representatives are also severely circumscribed by the operation of 'free' market forces. Socialists, by contrast, support a participatory or direct democracy, involving the ongoing involvement of all in day-to-day political debate, negotiation and decision-making. And such decision-making includes the formulation and implementation of economic policy.

Where delegation or representation is necessary, worker delegates or representatives are clearly mandated to represent the views of the workers involved, and report back to such workers on a regular basis. They are recallable and replaceable at any time if they are deemed to be failing in this regard.

In one model, that of Pat Devine, democratic planning 'takes the form of a political process of negotiated co-ordination, with decisions being made, directly or indirectly, by those who are affected by them'.[5]

> Broad economic parameters – covering such matters as the macro-economic division of resources between individual and collective consumption, social and economic investment, energy and transport policies, and environmental priorities – would be decided nationally – [and internationally] by [an] elected representative assembly on the basis of a set of alternative plans drafted by experts. But, within this framework, the bulk of economic decision-making would take place on a decentralised basis. Economic power would be vested in negotiated co-ordination bodies for individual production units and sectors on which would sit representatives of the work force, consumers, suppliers, relevant government bodies, and concerned interest groups…[6]

The neo-liberal critique focuses mainly upon the central socialist or communist preoccupation with equality of outcome, with an essentially equal distribution of consumption goods, social tasks and social power. The claim is that capitalism is a viable system in light of the basic selfishness of human nature, because it forces lazy, greedy people to contribute, faced with the stick of destitution and the carrot of wealth and power. It also allows freedom for the exercise of deep human desires in running private businesses.

Without any need for 'political' intervention, the market itself offers material incentives of consumption goods and social power (and disincentives of deprivation of such goods and power) to motivate socially useful productive effort and innovation. At the same time, there will always be some whose selfish inclinations lead them to

'break the rules'. There will therefore always be a need for some form of repressive criminal law to keep such people in check.

'True' socialism or communism, with guaranteed jobs and welfare on the one hand and no opportunity to run a business or accumulate private wealth on the other, removes both stick and carrot of market forces, leading to the victory of laziness, inertia and lack of initiative on the one hand and frustration on the other. These, in turn, lead, at best, to radically reduced productivity, and depletion of total social product, with living standards for all declining, at worst, to total social disintegration.

Socialists, so the argument goes, presuppose rather an angelic human nature; all people are kind, thoughtful, unselfish and ready to make huge sacrifices for their fellows. They have to do this, for without the incentives offered by capitalist markets, all that is left to motivate productive efforts is such love of others and self-sacrifice – or else draconian political threats of police/military repression. We all know that people are far from being angels. Nor can they ever be turned into angels (as some communists suggest) by any imaginable social engineering processes; greed and selfishness are in the genes.

Liberal democracy gives all people a meaningful input into the political process, without requiring them to waste time in detailed deliberations they are ill-equipped to effectively pursue. Liberal equality of opportunity allows individuals – and society – to make the best use of particular individual talents and skills.

In attempting to equally distribute socially productive tasks, communists would pay a terrible price in terms of the gross inefficiency of a system that failed to capitalise upon the different skills, talents and interests of different individuals. Instead, individuals suffer frustration and anger through not being allowed to exercise such skills and talents. And attempts to achieve a complete equality of social power, through universal participation in all political and economic decision-making, will inevitably turn into the grossest inequality and oppression.

Socialists argue that central planning can be a democratic process, with ongoing input of all concerned, but such ongoing input (such 'direct democracy') is quite impossible in practice for a large-scale and hugely complex industrial society: universal input would be just too difficult, inefficient and time consuming; there is no way everyone could meaningfully contribute to the process or be considered in the outcome. A small clique will inevitably take control of the planning process, and thereby gain totalitarian power over society as a whole,

and their selfish human nature will ensure that production is thereafter geared to satisfying their own needs and wants, rather than those of the rest of society.

Given the corruption of such a leadership by total power, their burgeoning demands on the rest of the population, the absence of market discipline, and reward for productive effort, and the absence of market efficiency to create sufficient output to allow for a sharing of social benefits between workers and planners, the central planners will increasingly resort to political sanctions to motivate productive involvement (and squash individual initiatives in attempting to re-establish free market relations), leading to the rule of repression and terror.

Such a 'theoretical' argument is taken to be solidly supported by the available empirical evidence. Communism in Russia started out down a path of democratic planning to satisfy the needs of all, but turned into the nightmare of Stalinist despotism, with gulags, food queues, totalitarian one-party politics, show trials, mass murder, disintegrating nuclear reactors, dissidents confined to mental hospitals, striking farmers, black markets and clunky cars.

Social liberalism

Social liberalism (also known as welfare liberalism or social democracy) is frequently seen as occupying an intermediate position between socialism and right liberalism, trying to achieve the 'best of both worlds'. Social liberals try to retain the equality of reward and incentives, 'efficiency' in allocation and scope for 'free' contractual negotiation – offered by free market relations, as the driving force of economic growth – while also prioritising a significant degree of equality of outcome and opportunity through business regulation, redistributive taxation and comprehensive public welfare provision.

Social liberalism still prioritises individual rights of 'free choice' (of life goals) and reward for effort and productive contribution, as primary values. It allows a substantial private sector where such incentives and initiatives drive innovation and economic growth. But it recognises the – utilitarian – public responsibility to compensate for undeserved disadvantage, in pursuit of genuine equality of opportunity. Whereas neo-liberals tend to attribute all or most disadvantage and deviance to 'free choice' (to pursue a life of crime or ill health) and (lazy) 'human nature', social liberals appeal, rather, to ideas of social

determinism. In particular, deviance, crime and physical and mental health problems (including drug use), along with much poverty, are traced back to less than optimum conditions of early social life and situation.

Whereas neo-liberals tend to assume that individuals are always free to make meaningful life choices, unless subject to direct physical coercion, social liberals focus much more upon the social-situational obstacles to, and restrictions upon, 'free choice'. They focus upon ways to reduce such obstacles and restrictions so as to increase the scope for genuinely free and informed decision-making.

This is the basic rationale for more or less comprehensive public welfare provision. Market forces alone do not guarantee genuine equality of opportunity or of reward. They do not, alone, maximise social welfare. The state needs to step in to try to create a more level playing field, through ongoing redistribution and market regulation.

Social liberals recognise that free market forces produce increasing inequality, both through the creation of ever bigger and more powerful monopolies and oligopolies and through periodic crises of mass unemployment. They are ambivalent about private oligopolies, recognising their power to achieve economies of scale, their power in the world market and capacity to boost growth and government tax revenue at home. They seek to regulate the operation of such corporations, moderating the pursuit of profit with consideration of broader social goals.

The social-liberal agenda includes public ownership of 'natural monopolies', with substantial efficiency gains from co-ordinated management and avoidance of wasteful duplication of infrastructure (in supply of electricity, railways, telecommunications, etc), and other key productive resources. Monopoly profits from some sectors can then subsidise necessary public services in others, as well as contributing to welfare spending – on health, education, social security and legal aid. Public services are also funded through progressive taxes upon income or wealth (including death duties and taxes on dividends and capital gains) supplemented by indirect taxes on luxury purchases.

Ongoing state fiscal and monetary intervention aims to stabilise markets in pursuit of full employment. In particular, deficit spending (including spending financed through sale of treasury bonds, along with tax cuts for poorer people and reduced interest rates) in the slump phase of the business cycle increases aggregate demand, with aggregate supply rising to meet such demand and bringing unem-

ployed resources – including people – back into employment.

In the boom phase of the cycle, increased company taxes (and interest rates) reduce the pressures towards overproduction and recoup public monies spent in the slump. Permanent high levels of employment provide (income and company) taxation to further support government welfare spending.

Such high levels of employment increase the power of organised labour in collective bargaining for wages and conditions. Strong trade unions and effective legal regulation of business can improve conditions of health and safety at work, and trade practices legislation can protect consumers from dangerous and defective purchases. Comprehensive no-fault insurance and unemployment benefits protect workers from income loss through days off work and medical expenses. Healthy, happy and empowered workers are, in turn, more productive workers, contributing to increased output to sustain investment and economic growth.

Complete freedom of international trade and investment threatens such market regulation, with market stimulation undermined by increased imports and sales of the local currency in international money markets, leading to devaluation and increasing trade deficits. Similarly, increasing company taxation is undermined by international capital mobility. To counter such developments, social liberals have advocated capital controls, Tobin taxes (on international currency transactions), protection of local industries and moves towards increasing local – or regional – self-reliance and autarchy.

Social liberals' commitment to freedom of speech is also tempered by utilitarian concerns and by concerns for rights deemed more pressing in the circumstances. They support such developments as criminalisation of racial vilification, restrictions upon political advertising and legally enforceable requirements of balance and fairness in media reporting and comment. They support limits upon political campaign funding, with all candidates ideally given equal 'exposure' to voters, with public subsidies to the less well off.

Whereas neo-liberals are unwilling to criminalise anything other than specific threats or harms to individuals' lives and property, social liberals defend criminalisation of what they see as threats to 'the common good' or the community as a whole or to more vulnerable groups within the community. In particular, they defend (criminal) legal steps to enhance environmental protection, public safety, food safety and other such 'public goods'.

Social liberals support representative parliamentary democracy, but they are concerned with the potential 'tyranny of the majority' and homogenisation of politics inherent in the Westminster 'first past the post' system favoured by neo-liberals. In face of these sorts of problems, social liberals tend to favour alternative systems of proportional representation, in which political groups are represented in legislative and executive bodies in proportion to the votes cast for them by the electorate as a whole. Here 'all votes end up counting towards the final result and are not wasted'.

> This allows people to vote more with their 'conscience' and according to their desires rather than being put in a position of having to choose tactically the lesser of evils to ensure their votes will count.[7]

Here, effective action depends upon negotiated consensus, co-operation and multiparty coalition, reconciling the differing views of different sections of the electorate, rather than the bully-boy tactics of individual 'strong leaders'.

The neo-liberal critique of social liberalism is straightforward. Apart from being an attack on basic property rights, redistributive taxation radically undermines equality of reward. It removes the incentives that drive entrepreneurial skills and innovations, while welfare rewards the lazy and unproductive. The result is declining social wealth and declining living standards for all.

The inevitably inflationary consequences of the deficit spending policies employed by social liberals in power to try to iron out the booms and slumps of the business cycle exacerbate these problems. Lazy debtors find their debts evaporating, while those who have worked hard to accumulate assets find the value of their assets reduced. Inflation reduces investment due to uncertainty about future terms of exchange, with everyone losing out as a consequence.

Neoclassical theory argues that free markets 'naturally' tend to full employment equilibria, with resources all fully employed and the system operating at peak capacity. This does not actually mean that everyone is employed, but rather that everyone who wants to be employed at the wage levels businesses offer is employed, with the others 'voluntarily' unemployed. In this situation, any government attempt to bring these unemployed into employment through pumping more money into the economy – thereby increasing purchasing power – will only lead to increases in the price of all goods, that is, inflation. Because the system is already at or close to peak

output, with no idle capital available to increase output, consumers with extra purchasing power will try to outbid one another for the limited goods actually available, and producers will similarly compete for the limited resources available to satisfy the increased demand, leading to increased prices across the board.

Those who are now prepared to take on employment at the new, increased wage levels soon find their real wages eaten away by such inflation, as employers pass on their increased production costs. They therefore leave the work force once again, and those who are still employed are now factoring runaway inflation into their future wage demands, thereby ensuring that such inflation does indeed come about. All are ultimately worse off than before the government intervention.

Social liberals' belief in social determinism – of deviance – leaves them open to neo-liberal criticisms that they are soft on crime, and inconsistent in prioritising 'free choice' and individual responsibility in some areas but not others. So are social liberals inconsistent in failing to allow 'free speech' in all areas. Everyone should be free to present their case, however appalling it might seem to some. Then everyone else is free to make up their own mind without paternalistic government interference.

If particular media magnates gain control of substantial channels of communication (through free market transactions), this will generally mean that they are giving consumers what they want. Such proprietors should be free to present whatever views they want, with consumers perfectly free to stop buying their papers or watching their TV channels if they choose to do so. Again, there is no justification for paternalistic government interference.

Similarly social liberals' commitment to welfare leads to neo-liberal claims that they contribute to the creation of 'welfare cultures' that crush effort, initiative and self-esteem, and transmit dependence from one generation to another. At the same time, issues of aging populations, increasing costs of medical technologies and, above all, globalisation, render such social welfare provisions quite unsustainable in the longer term.

According to neo-liberals, in a world of freely mobile big capital, free international trade and free currency markets, the high company taxation, corporate regulation and high inflation necessary to pay for the increasing costs of welfare will all be severely punished. High levels of corporate taxation will drive investment overseas to lower-taxed,

less regulated regions. High inflation will drive mass sales of the country's currency in international currency markets, leading to radical falls in its value against other currencies. This latter development might be expected to increase exports, but remaining highly taxed and regulated home-based industries will be radically disadvantaged in competition with overseas producers, with businesses and workers forced to pay more for necessary imports. So the upshot will be burgeoning unemployment, trade deficits and indebtedness.

In such an open world economy, money pumped into the local economy by the government can contribute to job creation overseas, rather than in the domestic economy, through purchase of imports and overseas investment. So too can it contribute to trade deficits and currency devaluation.

Social-liberal attempts to address these issues inevitably take the form of increasingly repressive state regulation: wage controls, capital controls, trade barriers, which will crush local initiative and lead to increasing national isolation and poverty, with loss of the benefits of free world trade and investment. Prices will rise, standards will fall. These developments threaten to completely undermine the national economy.

In the neo-liberal view, proportional representation is an obstacle to political and economic stability and progress. It fosters endless indecisive debate, rather than the swift and decisive action of a strong political leadership, willing and able to take difficult but necessary decisions.

Discussion topics

1. What is neo-liberalism? How does it relate to law?
2. What is social liberalism? How does it relate to law?
3. Critically evaluate the neo-liberal critique of socialism.

Additional resources

S. Mann, *Economics, Business Ethics and Law*, Lawbook Co, Sydney, 2003, ch 4.

Chapter 12

MARXISM AND LAW

This chapter provides a basic introduction to Marxist conceptions of the state and law. As noted in the previous chapter, Marxist ideas are radically different from those of both neo-liberals and social liberals.

1 Where western law asserts the sanctity of private property, freedom of contract and the 'rule of law' itself, as supposed guarantors of liberty and formal equality, Marxists argue that these doctrines inherently produce economic and social inequality.

2 While western law enforces the stability of the nuclear family as an economic unit, Marxists call for genuine freedom of choice in undertaking and leaving marriage and gender equality in family and social relations.

3 Whereas western law declares miscreants punishable because of their alleged personality defects, Marxists regard 'crime' primarily as a product of social inequity and, accordingly, seek to replace 'punishment' with social improvement, education and other remedial measures.

4 Western jurists insist that law is an organic and indispensable method of governing society, essential to combat or curb the alleged deficiencies and aggressive tendencies of human nature. Marxist jurisprudence regards humanity as capable of rising to a higher social and moral level, given the right conditions. It views the state and law as legacies of exploitative class society and seeks

to create the social conditions for them to be supplanted by more participatory and democratic forms of administration.

But what is the relevance of this today? Has not communism failed? In the early 1990s, certain writers asserted that the demise of the Soviet Union and the Eastern European Stalinist regimes signalled the irrevocable triumph of the market over socialism and even the 'end of history', to use Francis Fukuyama's phrase.[1] It is necessary to review the historical record in order to assess these claims.

The early years after the 1917 Russian Revolution produced groundbreaking achievements in legal policy. In several spheres, Soviet approaches were the most progressive in the world. They included the transformation of family and sexual relations – the recognition of the rights to divorce, de facto marriage and abortion – and the decriminalisation of the official response to anti-social behaviour. Underpinning these initiatives were the broader abolition of private ownership of basic production and finance, as well as efforts to de-formalise and provide for popular participation in social administration.

However, these experiments were cut short by the severe difficulties of the civil war and New Economic Policy (NEP), followed by the Stalinist degeneration.

The adoption of the NEP in 1921 caused a shift back to legalism, particularly with regard to the protection of private property rights. After late 1923, with the ascendancy of Stalin and the doctrine of 'socialism in one country', a new atmosphere of 'corrections' and diatribes set in, accompanied by a strengthening of the repressive state apparatus. The classical Marxist perspective of the withering away of the state and law was ditched in favour of the entrenchment of a legal edifice, erected in the name of 'socialist legality'. In this sphere, as in others, Stalinism was a repudiation of Marxism, not a continuation of it.

These implications can be seen in the fate of Eugene Pashukanis, the best-known early Soviet jurist. Pashukanis' 1924 commodity theory of law initially became part of the regime's official doctrine. It helped reconcile the needs of the NEP, including the legal protection of private property rights, with the Marxist understanding of the withering away of the state. Despite various dismissive critiques in the West,[2] however, his theory provided some profound insights into the nature of law under capitalism.[3]

By the late 1920s, Pashukanis was under attack within the Soviet Union because he maintained, in keeping with authentic Marxism,

that the law and indeed the state apparatus of the Soviet Union would ultimately disappear with the construction of a genuinely communist society. He initially resisted the Stalinist notion that the law and the state itself had become organically 'socialist' and therefore occupied a permanent place in social organisation. By 1935, his views were incompatible with the Kremlin line, which was based on the wholly self-contradictory claim that socialism had been built; yet the 'dictatorship of the proletariat' had been simultaneously strengthened.

Marxist jurisprudence

None of the leading Marxists – Marx, Engels, Plekhanov, Lenin, Trotsky – attempted to set out a comprehensive model of society under socialism. They regarded such ventures as overly prescriptive, as well as premature and utopian. For them, socialism consisted of human self-emancipation and would be shaped by the actions and ideas of millions of working people, tempered by the concrete historical and international circumstances that prevailed. The classical Marxists were even less inclined to provide a detailed blueprint for the role of law and the state machinery in the transition from the overthrow of capitalism to socialism and then communism. They regarded law's role as being fundamentally bound up with, and, in the final analysis, dependent upon, the development of humanity's economic capacities and social wellbeing.

Nevertheless, while Marx and Engels did not write systematic expositions on legal theory, many of their works examined the role of law in society. They provided a definite framework of analysis and orientation, as well as basic principles, which initially guided the early leadership of the Soviet Union but were later betrayed under Stalinism.

The two fundamental underlying Marxist conceptions are that, in general, all forms of law and the state were in the end derived from the development of the productive and hence cultural level of human society and, that law and the state would wither away in the process of arriving at a genuinely communist society. That is, the need for formal, bureaucratic and repressive instruments of rule would disappear with the creation of a bountiful, egalitarian and democratic world.

The starting point for understanding this historical materialist view is Marx's 1859 Preface to *A Contribution to the Critique of Political Economy*, where he tentatively described the following propositions, derived from years of research and experience, as 'a guiding thread for my studies':

> In the social production of their life, men enter into definite relations that are indispensable and independent of their will, relations of production which correspond to a definite stage of development of their material productive forces. The sum total of these relations of production constitutes the economic structure of society, the real foundation, on which rises a legal and political superstructure and to which correspond definite forms of social consciousness. The mode of production of material life conditions the social, political and intellectual life process in general. It is not the consciousness of men that determines their being, but, on the contrary, their social being that determines their social consciousness.
>
> At a certain stage of their development, the material productive forces of society come in conflict with the existing relations of production, or – what is but a legal expression for the same thing – with the property relations within which they have been at work hitherto. From forms of development of the productive forces these relations turn into their fetters. Then begins an epoch of social revolution. With the change of the economic foundation, the entire immense superstructure is more or less rapidly transformed.
>
> In considering such transformations, a distinction should always be made between the material transformation of the economic conditions of production, which can be determined with the precision of natural science, and the legal, political, religious, aesthetic or philosophical – in short, ideological forms in which men become conscious of this conflict and fight it out.[4]

Three themes can be discerned in this seminal passage. The first is that law, like other aspects of the political superstructure, arises from definite relations of production and the forms of social consciousness forged by those relations. The second is that those relations are not static but are inevitably shattered by the further development of technology and production itself, ultimately leading to social revolution. The third is that law is one of the ideological forms in which humanity becomes conscious of the underlying conflicts and 'fights them out'.

Essential propositions

Properly understood, the Marxist view of law includes a number of pivotal propositions. First, socialism means democracy and the withering away of the state, not the bureaucratic 'command economy' that subsequently emerged in Soviet Russia under Stalin.

Second, socialism cannot be achieved by seeking to reform the state machine of the old order. It requires a thoroughgoing popular revolu-

tion to establish a new kind of state, a genuinely democratic state (the dictatorship of the proletariat), as a transitional regime to create the ultimate conditions for a classless, stateless communist society.

Third, law is not inherent or organic to society; rather it arises out of conflicting interests in society and primarily reflects the interests of the ruling layers. Therefore, in a classless society, the legal form of social regulation will become redundant. This withering away of the state and law can and must begin as soon as the socialist revolution has successfully wrested power from the old ruling class.

Fourth, the relationship between law and socio-economic power is dialectical. Against crude materialism and class reductionism, Marxists explain that legal definitions and measures can, in some circumstances, exert a sharp influence on economic and social developments. In part, this arises from the mystified, ideological form in which law and legal theory present themselves.

Finally, the Marxist view of law rejects the notion that capitalism, based on private ownership of the means of production, is somehow natural while socialism is alien to human nature. Under capitalism, law also plays an ideological role in disguising social inequality, dulling consciousness of class divisions and reinforcing 'commodity fetishism'.

Socialism, democracy and the state

Marx argued that the development of a socialist society will not take place according to a series of prescriptions and rules laid down by an individual, a political party or a governmental authority. Rather, it will develop on the basis of the activity of the members of society who, for the first time in history, consciously regulate and control their own social organisation as part of their daily lives, free from the domination and prescriptions of either the 'free market' or a bureaucratic authority standing over them.

In Marx's view, the precondition for such a society is the development of the social productivity of labour to such a point that the vast bulk of humanity does not have to spend the greater portion of the day merely trying to obtain the resources to live. The overturn of capitalist rule would not see the overnight abolition of the market. The price mechanism would still be needed for a period as a guide in the provision of information regarding the relative costs of alternative production methods. But increasingly it would be made subordinate

to and eventually replaced by the conscious regulation of the economy according to a plan, decided on, checked and altered to meet changing circumstances through the involvement of workers and the population as a whole in process of economic decision-making.

The emergence of the Stalinist bureaucracy in the early 1920s, and its complete usurpation, by 1927, of political power meant that genuine socialist, that is, democratic, planning could never be carried out in the Soviet Union. Such democratic input would have immediately threatened the privileged social position of the bureaucracy and its monopoly of political power.

Democratic participation was an essential prerequisite and ongoing requirement for the harmonious development of a genuinely socialist economy and the all-round growth of productive output, as well as social emancipation. This imperative has enormous implications for law, being a central component of the need to de-legalise social life as far as possible and facilitate the withering away of the state.

The transition to communism

The dictatorship of the proletariat, in the writings of Marx and Engels, means the temporary and emergency political rule of the working class, as the first stage in the transition to a classless, stateless society. This political rule must include the control by the associated producers – the working class, which constitutes the overwhelming majority of society – of the productive forces they themselves have created. In other words, the dictatorship of the proletariat means from the outset the establishment of genuine democracy, with the majority of the population exercising economic power.

The term 'dictatorship of the proletariat' as used by Marx and Engels does not mean tyranny or absolutism or rule by a single individual, a minority or even a single party but political rule exercised by the majority of the population. This is clear from their analysis of the Paris Commune of 1871, which ruled Paris for a period of 72 days before being militarily crushed. In his 1891 Introduction to the reissue of Marx's analysis of the Commune in *The Civil War in France*, Engels explained that the Commune, which was the first attempt at establishing the dictatorship of the proletariat, began with the 'shattering of the former state power and its replacement by a new and truly democratic one'.[5]

In the *Critique of the Gotha Program*, Marx distinguished between the two stages of socialism. In the first, it would be impossible, given

the economic, intellectual and moral birthmarks of the old capitalist order from whose womb socialist society emerged, to go beyond the 'narrow horizon of bourgeois right', by which he meant the formal legal equality that invariably masks social inequality. 'Law can never stand higher than the economic order and the cultural development of society conditioned by it', Marx wrote.[6] That is, the law would inherently reflect the fact that society could not provide a plentiful and satisfying life for all. Only after individuals were no longer enslaved by others, labour had become a meaningful and enjoyable pursuit rather than a burden, and the productive forces had increased abundantly, would the communist ideal be realised.

Trotsky defended this underlying conception in *The Revolution Betrayed*, his analysis of the degeneration of the Soviet Union:

> The material premise of communism should be so high a development of the productive forces that productive labour, having ceased to be a burden, will not require any goad, and the distribution of life's goods, existing in continual abundance, will not demand – as it does not now in any well-off family or 'decent' boarding-house – any control except that of education, habit and social opinion.[7]

Central to this view, as first expounded by Marx and Engels in the *Communist Manifesto* and later by Lenin in *The State and Revolution*, was that the state and law must begin to fade away as soon as the dictatorship of the proletariat was established. That is, inherent in the seizure of political power and the establishment of a workers' state was the creation of a unique kind of government that would immediately begin to transfer society's administration into the hands of the population at large.

Engels returned to this theme in his 1891 Introduction to Marx's *The Civil War in France*, which described the formation and suppression of the Paris Commune. Engels contrasted the Commune to all previous revolutions, which had replaced one oppressive state by another.

> From the very outset, the Commune was compelled to recognise that the working class, once come to power, could not go on managing with the old state machine; that in order not to lose again its only just conquered supremacy, this working class must, on the one hand, do away with all the old repressive machinery previously used against itself, and, on the other, safeguard itself against its own deputies and officials, by declaring them all, without exception, subject to recall at any time.[8]

Apart from the right of recall, Engels reviewed three other measures taken to prevent 'careerism': election to all posts – administrative, judicial and educational; restriction of the wages of all officials, high and low, to those paid to workers; and binding mandates for delegates to representative bodies.

Interaction between law and social structure

In many western academic writings, Marx and Engels are presented as mechanical economic determinists. This somewhat simplifies their analysis. They were determinists in the following sense. For them, the driving forces of all economic, political and social life are the contradictions in material and economic life. Essentially, these contradictions arise from the conflict between the social forces of production and the relations of production – the class and property relations of society – within which those productive forces have hitherto developed.

More specifically, the development of capitalist economic relations shaped the content and structure of law in many ways. The most fundamental relate to the core concepts of private property and contract. Both required an essential break with feudal relations, based on communal and feudal property, fixed status and personal allegiance. Capitalism, as an expansionary economic system, demanded the unfettered accumulation of capital based on the private ownership of the means of production.

Marx and Engels concluded that the ultimate driving forces of all economic, political and social life are the contradictions in material and economic life. This analysis is far from passive, lifeless and mechanical. While the decisive factors shaping law are economic relations, the legal system remains one of the arenas within which the class struggle is fought out. As Engels pointed out in his 1890 letter to Conrad Schmidt, Marx's section on the working day in *Capital* shows that legislation can have a 'drastic effect' on social conditions and the class struggle.[9]

This conflict is not automatically reflected in legal doctrines but refracted through the need to elaborate legal principles that have the appearance of internal coherence and universality and to continually adjust those doctrines to meet changing economic circumstances. On law, as other social phenomena, Marx and Engels demonstrated the dialectical interaction between the economic base of society and the ideological superstructure.

In a letter to J Bloch, Engels emphasised that the economic situation is the 'ultimately determining factor in history' but

> the various elements of the superstructure – political forms of the class struggle and its results, such as constitutions established by the victorious class after a successful battle, etc, juridical forms, and especially the reflections of all these real struggles in the brains of the participants, political, legal, philosophical theories, religious views and their further development into systems of dogmas – also exercise their influence upon the course of the historical struggles and in many cases determine their *form* in particular [italics in original].[10]

This analysis was also dynamic in relation to the continual contradictions produced by the further development of the productive forces and new forms of property rights. Further contradictions arose constantly from the ideological role of law: from the need of any modern ruling class in the epoch of mass politics to present its political order as just and impartial. In his letter to Conrad Schmidt, Engels stated:

> In a modern state, law must not only correspond to the general economic condition and be its expression, but must also be an *internally coherent* expression which does not, owing to internal conflicts, contradict itself. And in order to achieve this, the faithful reflection of economic conditions suffers increasingly. All the more, so the more rarely it happens that a code of law is the blunt, unmitigated, unadulterated expression of the domination of a class – this in itself would offend the 'conception of right'.[11]

Law and ideology

While Marx and Engels recognised that under capitalism, ideological factors could determine the form of legal development, the resulting process produced a mystification, by presenting economic interests as philosophical principles. In his 1890 letter to Schmidt, Engels wrote:

> Economic, political and other reflections are just like those in the human eye. They pass though a condensing lens and therefore appear upside down, standing on their heads. Only the nervous system, which would put them on their feet again for representation, is lacking … The reflection of economic relations as legal principles is necessarily also a topsy turvy one: it happens without the person who is acting being conscious of it; the jurist imagines that he is operation from *a priori* principles, whereas they are really only economic reflexes; so everything is upside down. And it seems to me obvious

> that this inversion, which, so long as it remains unrecognised, forms what we call ideological conception, reacts in its turn upon the economic basis and may, within certain limits, modify it.[12]

Precisely because law was a distorted reflection of economic reality, the distortion could, to the extent that the deformation went unrecognised, impact on the underlying economic relations. This view has a number of implications. In the first place, the mystified distortion served to legitimise exploitation. By reproducing in legal form the commodification of all relations, law presented these relations in an 'inverted' way, camouflaging their real content. This was not simply a conspiracy or confidence trick perpetrated by the ruling class, aided by legal theorists and lawyers. Because law was shaped by the objective requirements of the capitalist mode of production, it was organically shrouded in a distorted view of social relations. Bourgeois legal theorists were themselves trapped in an ideological inversion, mistakenly regarding their ideas as the source of jurisprudential development.

Engels elaborated on the role of ideological factors in several of the letters written in the final years of his life. In an 1893 letter to Franz Mehring, Engels sought to correct a 'mistake' that he and Marx may have made in focusing mainly on the basic economic facts while neglecting the ideological forms through which the economic factors were expressed.

> Ideology is a process accomplished by the so-called thinker consciously, indeed, but with a false consciousness. The real motives impelling him remain unknown to him; otherwise it would not be an ideological process at all. Hence, he imagines false or apparent motives. Because it is a process of thought, he derives both its form and its content from pure thought, either his own or that of his predecessors. He works with mere thought material which he accepts without examination as the product of thought; he does not investigate further for a more remote process independent of thought; indeed, its origin seems obvious to him, because, as all action is produced through the medium of thought, it also appears to him to be ultimately based upon thought.[13]

Law and human nature

The proponents of the free market and capitalist ownership of the means of production argue that socialism is unnatural and therefore doomed to failure because it violates the inherent drive in every human being towards the exclusive ownership of property. This conception is filled with unstated assumptions.

In the first chapter of *Capital*, in his analysis of commodity fetishism, Marx explained that one of the great difficulties in coming to an understanding of society is that it has already undergone a considerable development:

> Man's reflections on the forms of social life and consequently, also, his scientific analysis of those forms, take a course directly opposite to that of their actual historical development. He begins, *post festum*, with the results of the process of development ready to hand before him.[14]

In other words, analysis begins with categories and forms of thought already at hand, under conditions where the historical processes that gave rise to these forms is obscured from view. Hence these forms of thought are not understood as the product of historical processes, but seem to spring from the 'inner nature' of man himself. Take, for example, the question of interest. Nothing may seem more natural than that there should be a payment or interest charged on the use of money. Under capitalism, economic life would quickly grind to a halt if lending for profit ceased. Yet, for hundreds of years there were denunciations of usury and severe punishments inflicted for its practice. Moneylending was depicted as sinful, essentially because it threatened to undermine feudal relations, which were based on status, not money.

In capitalist society, the extraction of surplus labour does not take place through political means, but economically. That is, while there were a myriad of laws in feudal society, which spelt out the obligations of the peasant, there are no such laws under capitalism. There is no statute that compels the worker to sell his or her labour power to the owner of capital. He or she is forced to do so by the pressure of economic necessity. And that compulsion arises from the fact that, unlike the peasant or small producer in feudal society, the worker in capitalist society has been separated from the ownership of the means of production.

Therefore, the crucial question to be examined in the transition from feudalism to capitalism is how this transformation took place. That is, how it was that a class of free wage labourers emerged – free both from feudal obligations and from the means of production – with nothing to sell but their labour power. History shows that this transformation did not result from some innate human nature, but was the outcome of new forms of social organisation based on the market. Those who maintain that the emergence of capitalism is the result of

some inherent drive to own private property can never answer the question as to why the transition to capitalism took place between the 16th and 18th centuries, rather than earlier. Capitalism could only emerge once society's technology and productive capacity – for example, steam power – had developed to the point where large-scale manufacturing could arise.

One of the most important battles in the development of capitalism was the establishment of *exclusive* property rights, above all in land, over the common property rights that had played such a central role in the lives of the peasantry under feudalism. Far from expressing some inherent human characteristic, manifesting itself at a young age, this new form of property had to establish itself against the conception that land should be held in common and its fruits available to all. Locke, in particular, had to argue strenuously for the right to individual property, against the conception of common property and custodianship. Locke identified certain inalienable rights: the right to life, liberty and property. According to Locke, every man was the sole proprietor of his own person and capacities. His right to property derived from his right to enjoy the fruits of his own labour.

The theory that identifies freedom with private ownership is based on the claim that each individual has the natural right to the fruits of their own labour and that private property is the means through which this right is secured. But concentration of ownership and the separation of the mass of the population from the means of production with nothing to sell but their labour power to the owners of capital means that private property itself has long ago undergone a transformation. No longer is it a social mechanism through which individuals secure the fruits of their *own labour*; it is rather the mechanism through which capital secures the fruits of *other people's labour* in the form of profit.

Leon Trotsky's observations

Leon Trotsky was one of the foremost leaders of the October 1917 Revolution and of the fight against Stalin's bureaucratisation of Soviet society. Trotsky defended the October 1917 Revolution and the initial actions of the Bolsheviks, including the seizure of power, the dissolution of the Constituent Assembly and the banning of parties that took up arms against the revolution and other measures taken during the civil war of 1919–21. In *Terrorism and Communism – A Reply to Karl*

Kautsky, Trotsky responded to Kautsky, previously a major figure in the Marxist movement, who accused the Bolsheviks of proceeding undemocratically. Members of the Austro-Marxism school, who, in some instances, claimed to be Marxist legal theorists, joined Kautsky's denunciation of the revolution. They included Karl Renner, Otto Bauer, Max Adler, Rudolf Hilferding and Friedrich Adler.

Trotsky insisted that the Soviet revolution was far more democratic than the parliamentary apparatus defended by Kautsky. He pointed to the innate fraud of capitalist democracy, which leaves the economic power and control over the state apparatus in the grip of a ruling elite, arguing that it gives the working masses no other way but revolution to take charge of society.

Trotsky also examined democracy from a theoretical and historical standpoint. He pointed to the degeneration of the democratic conception in the hands of the capitalist class and its jurisprudential theorists.

> As a battle cry against feudalism, the demand for democracy had a progressive character. As time went on, however, the metaphysics of natural law (the theory of formal democracy) began to show its reactionary side – the establishment of an ideal standard to control the real demands of the labouring masses and the revolutionary parties … Natural law, which developed into the theory of democracy, said to the worker: 'all men are equal before the law, independently of their origin, their property, and their position; every man has an equal right in determining the fate of the people'. This ideal criterion revolutionised the consciousness of the masses in so far as it was a condemnation of absolutism, aristocratic privilege, and the property qualification. But the longer it went on, the more it sent the consciousness to sleep, legalising poverty, slavery and degradation: for how could one revolt against slavery when every man has an equal right in determining the fate of the nation? … In the real conditions of life, in the economic process, in social relations, in their way of life, people became more and more unequal; dazzling luxury was accumulated at one pole, poverty and hopelessness at the other. But in the sphere of the legal edifice of the state, these glaring contradictions disappeared, and there penetrated only unsubstantial legal shadows.[15]

Trotsky related the need for the dictatorship of the proletariat in the transition to communism, to the political and economic tasks involved in overthrowing capitalism. He argued that genuine socialism and communism were impossible to achieve without the free and creative involvement of all people. The bureaucratic police state erected by Stalin was not only an affront to socialist democracy but also a suffo-

cating barrier to the development of the productive and cultural capacities of society.

> The socialistic economy must be directed to ensuring the satisfaction of every possible human need. Such a problem it is impossible to solve by way of commands only. The greater the scale of the productive forces, the more involved the technique; the more complex the needs, then the more indispensable is a wide and free creative initiative of the organised producers and consumers. The socialist culture implies the utmost development of the human personality. Progress along this path is made possible not through a standardised cringing before irresponsible 'leaders', but only through a fully conscious and critical participation by all in a socialistic creative activity. The youthful generations stand in need of independence, which is wholly consistent with a firm leadership but rules out any police regimentation. Thus the bureaucratic system in crushing the Soviets and the party is coming ever more clearly into opposition with the basic needs of economic and cultural development.[16]

In line with Marx and Engels, Trotsky's emphasis was on the self-liberation of the entire population. This was the essence of communism. Therefore, the task of the Soviet state was to encourage, not stifle, the maximum degree of conscious, well-informed and independent participation in political and administrative affairs.

Discussion topics

1. What is Marxist theory? How does it relate to law?
2. What is the role of law in a socialist society?
3. Can law 'whither away'?
4. Is law, in the final analysis, an instrument of class rule?

Additional resources

M. Cain and A. Hunt, *Marx and Engels on Law*, Academic Press, London, 1979.

F. Engels, *Ludwig Feuerbach and the End of German Classical Philosophy*, Progress Publishers, Moscow, 1978.

F. Engels, *The Origin of the Family, Private Property and the State*, International Publishers, New York, 1942.

M. Head, 'The Passionate Legal Debates of the Early Years of the Russian Revolution' (2001) 14 *Canadian Journal of Law and Jurisprudence* 3–27.

M. Head, 'The Rise and Fall of a Soviet Jurist: Evgeny Pashukanis and Stalinism' (2004) 17 *Canadian Journal of Law and Jurisprudence* (forthcoming).

V. Lenin, *The State and Revolution*, Progress Publishers, Moscow 1970.

C. Macpherson, *The Political Theory of Possessive Individualism: Hobbes to Locke,* Oxford, Clarendon Press, 1962.

K. Marx, *The Civil War in France*, Progress Publishers, Moscow, 1948.

K. Marx, *A Contribution to the Critique of Political Economy*, Progress Publishers, Moscow, 1977.

K. Marx, *Critique of the Gotha Program*, International Publishers, New York, 1970.

P. Phillips, *Marx and Engels on Law and Laws*, Martin Robertson, Oxford, 1980.

L. Trotsky, *The Revolution Betrayed. What is the Soviet Union and Where is it Going?*, Pathfinder Press, New York, 1972.

L. Trotsky, *Terrorism and Communism*, New Park Publications, London, 1975.

Chapter 13

EFFICIENCY AND THE MARKET

Is the market efficient?

As noted in chapter 10, one way of looking at law and justice, which has come to be called *the economic analysis of law* or *law and economics*, has become influential in recent years. As Freeman says in *Lloyd's Introduction to Jurisprudence*: '[I]n the last 40 years, at least in the United States, [this approach] has come to dominate thinking about law, and not jut in the more obvious commercial areas.'[1] This 'economic analysis of law' is an increasingly large and complex body of thought. Here, we focus only upon certain key considerations underpinning such economic analysis, and upon some relevant real-world economic issues.

Perhaps the fundamental idea of the economic analysts is that the development of modern economics has provided objective measures of *efficiency* that can be used to guide legislation and judicial decision-making as well as strictly 'economic' decisions and actions. Insofar as inefficiency is equated with wasted resources and wasted effort – productive resources that could be used to make things people want left unused, and more effort expended in producing such useful things than is necessary – it is assumed that we all have a 'common' interest in efficiency maximisation. As Hahnel says: 'as long as resources are scarce relative to human needs and socially useful labour is burdensome … efficiency is preferable to wastefulness'.[2]

According to both the *classical economics* of Adam Smith, and the more recently developed *neoclassical approach of economics* (strongly favoured by neo-liberals and economic rationalists), 'free markets' are by far the best way of achieving such 'efficiency'. Proponents of the

economic analysis of law are supporters of (elements of) the classical and the neoclassical approach (as opposed to other approaches, such as those of Marx, Keynes, or the post-Keynesians) and they argue that the law, in support of efficiency, should aim to facilitate – rather than in any way impede – the creation or maintenance of free markets, and should generally aim to compliment or mimic market mechanisms in its decision-making processes.

This raises many issues. Here we focus upon three groups of questions:

1 Is such 'efficiency' (as understood by classical or neoclassical economists) a desirable or defensible goal of social policy?

2 Are 'free' markets (as understood under classical and neoclassical theories) optimally 'efficient' in the terms in which these theories understand 'efficiency'? Do these ideas make moral and practical sense in theory?

3 Are current markets – or could they be – 'free' in the neoclassical sense of the term? Do these ideas make sense in practice? And to the extent that that they are not 'free', what are the implications of this for efficiency?

The first question is the difficult one. It raises two issues: What is a viable and useful conception of efficiency? Efficiency is far from easy to define. And how do we reconcile genuine efficiency with other values? If efficiency conflicts with justice or democracy or equality (including equal satisfaction of basic human needs), it is far from clear that the latter should always be sacrificed in favour of the former.

Adam Smith

Adam Smith (1723–90) argued that self-interested pursuit of profit in a free, competitive market is the driving force ensuring effective matching of supply to changing – monetarily effective – demand in the short term, and general improvements in living standards for all in the longer term. In a free market, the sale price of goods in increasing demand rises above the 'natural price' determined by production costs plus a 'natural' rate of profit with the price of goods that are in decreasing demand falling below such a natural price. Such increased prices register in increased profits for producers, so that profit-maximising entrepreneurs are thereby motivated to shift their investment in line with increased demand.

Similarly, since profits can also be increased in the short term through productivity gains, achieved through specialisation (both within and between particular production processes, regions and nations) and technological innovation leading to reduced production costs and consequently increased profits through sale at the 'old' natural price, rational profit-maximisers are also motivated to pursue ever greater productivity gains, thereby continuously reducing the real (labour time) costs of commodities. While the first innovators increase their sales and profits, others are forced to follow suit (adopting the new work practices and technologies) in order to survive, and prices are competed down to the new (reduced) production costs, plus natural profit. This, in turn, allows the working population access to an increasing range of such goods. The excess profits of the innovators allow them to expand their operations through accumulation of such profits, thereby further increasing the scope for higher productivity, greater employment, markets and consumption.

Thus, as long as obsessively acquisitive people (entrepreneurs with access to productive resources) are also rational people, their efforts to pursue their own self-interest cannot help but serve the interests of others (the rest of the population who have to work for them). This is the famous 'invisible hand'. While Smith believed that wages would remain around subsistence level – with higher wages encouraging working class population growth, intensified competition for jobs, and consequent falls in wage levels – he saw the continued growth of the system through capital accumulation in pursuit of productivity gains as maintaining demand for labour. Likewise, the declining real costs of commodities would allow for improved levels of 'subsistence', with an increasing range of goods included in the real subsistence income of the working population.

Neoclassical ideas

At the heart of neoclassical thinking are the so-called *laws of supply and demand*. The number of units of a good that suppliers will offer for sale will increase as the price they receive for the good increases, and consumers will offer to buy a declining number of units of a good as its price increases.

Key concepts here are those of increasing marginal costs and declining marginal utility. Because, above a certain level of output, sellers (supposedly) produce under conditions of rising cost, with each

new unit costing more to produce than the previous one, they require increased compensation (per unit) to motivate such increased production. And less efficient – higher-cost – producers can afford to produce a good as its sale price goes up. Similarly, because the usefulness to consumers of a further unit of a good decreases the more of that unit they already have, they will only buy further units if the price comes down. And poorer people can afford to buy a good at a cheaper price.

In a situation of *perfect competition*, where many sellers provide identical products with none able to affect the price of their product and many buyers with none able to affect the price they pay, and where all buyers and sellers have perfect information about all relevant products and their prices, all units of a good will sell at the same price and the uniform market price will adjust until the number of units buyers want to buy is equal to the number of units sellers want to sell. Supply and demand curves then intersect at what is called an *equilibrium price* and *equilibrium quantity*.

When the market price is above the equilibrium price those not selling their products (due to insufficient demand) will reduce their prices. When it is below, unsuccessful buyers will push it up by increasing the prices they offer for goods.

> So according to the … law of supply and demand, the only stable price will be the equilibrium price [for a particular sort of good] because self-interested behaviour of frustrated sellers or buyers will lead to changes in price under conditions of both excess supply and excess demand, and only at the equilibrium price is there neither excess supply nor excess demand.[3]

So too, does such a 'free' market – with free mobility of capital – respond quickly and effectively to changing consumer tastes and changing production costs due to technological advance. Increased demand for a product – due to changes in wage levels, for example – registers in the demand curve shifting to the right as consumers demand more of the product at every price. This drives up its price until the excess demand is eliminated. The new higher price for the product also leads to increased production to meet the increased demand, drawing resources away from areas of reduced demand and correspondingly reduced price. Similarly, reduced production costs will shift the supply curve to the right because of the increased quantity of the good that can be produced at any given price, leading to a lowering of prices till the excess supply is eliminated. Consumers buy more of the good at the reduced price, thereby taking advantage of the new technology.

Gross domestic product

It is difficult to relate these ideas to the visible workings of a national economy. Frequently, neo-liberal theorists and politicians simply focus upon gross domestic product (GDP) – the total monetary value of goods and services produced for the year in the territory in question – as indication or measure of overall economic efficiency, with increasing GDP indicating increasing efficiency. The assumption is that greater productivity in all or most areas produces greater profits that are invested to expand the scope and scale of production, leading to higher levels of consumption.

There are dangers of equating GDP with social welfare, or as a desirable or essential goal of social policy, including legal decision-making. The major problems lie in (a) the failure to consider issues of wealth and income distribution and (b) the failure to distinguish socially and environmentally useful production from destruction and waste. As Frank Stilwell says:

> Conventional GDP measures are … neutral as between different patterns of distribution. It matters not whether the additional goods and services being produced are consumed by just one individual or equally shared by all … But if … the principle of diminishing marginal utility applies to income, we cannot reasonably assume that the additional dollar's worth of consumption by a rich person adds as much to their economic wellbeing as the same dollar of consumption enjoyed by a poor person.[4] … [At the same time] when it comes to GDP calculations, the human costs of producing the goods and services is not set against the market value of those items … is means that it does not matter if some people are working very long hours while others are unemployed. It is only the aggregate that counts.[5]

In the United States, between 1977 and 1989, in a period of steadily increasing GDP, the average income of the top 1% of families rose by 78% and that of the bottom 20% decreased by 10.4%. Poorer people were working increasingly long hours for less.[6]

As Stilwell points out, the economists' concept of 'consumer sovereignty' serves as an excuse for them avoiding the task of distinguishing good, and socially valuable, production from bad.

> If a consumer chooses to spend his/her income on tobacco products, on a gas-guzzling automobile … on a noisy [polluting] jet-ski, that adds just as much to GDP as comparable expenditures on health, education and cultural enrichment. Guns count equally with butter.[7] … some forms of expenditure which count towards GDP are more con-

> cerned with damage repair than genuine economic progress. Health services provided to treat cancer resulting from excessive tobacco consumption add to GDP …[8]

Moreover, the large-scale destruction of the lifestyle and social infrastructure of subsistence farmers and foragers, with the creation of mass unemployment, destitution and beggary, along with dollar-a-day wages in dangerous sweatshops for a tiny minority – as continues to happen around the third world – counts as increasing GDP and increasing social welfare from an orthodox economic perspective.

Pareto optimality

In response to criticisms that they fail to consider the issue of distribution, neo-liberal theorists have made much of the concept of *Pareto optimality*. A Pareto optimal outcome is one where it's impossible to make anyone better off without making someone else worse off. A *Pareto improvement* involves making someone better off without making anyone else worse off – in monetary terms. Since neo-liberals argue that all are free to make their own decisions about participation in market transactions, and no-one will freely enter in a market exchange that makes them worse off, so a free market economy, by definition, moves from Pareto improvement to Pareto improvement

In fact, Pareto optimality is no better than GDP as a measure of efficiency or social welfare. In a society of radical inequality it is generally not possible to improve the lot of the worse off without some redistribution of the resources of the wealthy, through, for example, progressive taxation to fund social welfare and low-income supplementation. Even given the declining marginal utility of wealth, the very wealthy will still claim they are losing out, and can appeal to Pareto to prevent even the smallest of steps toward greater equality.

Pareto also appears to justify a step from 10 to 1 wealth and/or income distribution to 1000 to 1 (presumably through significant economic 'growth'). But, in fact, there is substantial empirical evidence that it is the extent of inequality, rather than the (absolute) level of wealth or income at the bottom, that is most significant in determining the quality of life for those at the bottom, as measured in terms of general physical and mental health, life expectancy, security, safety and wellbeing. (See the references to Wilkinson's work in earlier chapters.)

As we have seen, Rawls argues that the only excuse for increasing departures from equality is if they are necessary to improve the

condition of the worst off. This is at least an improvement on Pareto. But it too could be used to (attempt to) justify a move from 10 to 1 to 1000 to 2, perhaps through 'trickle down'. In fact, there is no reason to suppose that any increase at the top is necessary to improve things at the bottom, or that any increases at the top are actually likely to do so. On the contrary, historical evidence suggests that increasing inequality contributes to increasing absolute poverty. As Self points out:

> the market liberalisation policies pursued in Britain and the United States since 1980 and in Australia and New Zealand rather later, have resulted in very large increases in economic inequality and in the growth of absolute poverty, which contrasts with the trend towards less inequality and poverty in the previous decades of stronger government intervention in the market. In Britain the income of the richest one-fifth … rose from four to seven times the poorest fifth between 1977 and 1991, in the United States the same ratio … rose from 7.5 in 1969 to 11 in 1992. Change in the positions of the top rich and the bottom poor were still more startling, with the top tenth in Britain gaining almost 60% more income and the poorest losing almost 29% between 1979 and 1991. This rise in absolute poverty, which also occurred in the United States, is especially significant because it shows not *trickle down* but trickle up.[9]

It has frequently been argued that Pareto optimality has no implications for real-world policy decisions, which always make some better off and some worse off. If only Pareto improvements are allowed there will be no policies.

Cost benefit analysis

In light of these considerations, some neoclassical defenders of economic efficiency as a major social welfare goal focus rather upon *cost benefit analysis* as a seemingly more useful and defensible basis for defining such efficiency. The basic idea here is to compare the monetary value of the 'social costs' and 'social benefits' of any particular productive undertaking, with such an undertaking judged to be 'efficient' and to contribute to general social welfare to the extent of the monetary gain in benefits as against costs. For example, as Self explains, the cost benefit appraisal of road projects has, in the past, set against the monetary cost of road construction 'the expected reduction in accident costs and the time savings to drivers', with the latter valued 'on the basis of the average hourly earnings for drivers on business travel'.[10]

One reason for the appeal of this idea to neoclassical theorists is the way in which the basic principles of such analysis can be equated with the basic explanatory concepts and models of neoclassical microeconomic theory, as outlined earlier. The interaction of marginal social cost and marginal social benefit curves takes the place of the interaction of (market) supply and demand curves.

The idea is that if we plot the full social cost of producing increasing amounts of some particular good (cost on the y axis, amount on the x) we end up with a curve rising up from left to right; the social costs per unit increase due to the need to withdraw resources from other areas, increasing pollution, etc. Similarly, when we plot the full social benefits of the availability of each additional unit (benefit units on the y axis, amount on the x), we get a curve falling from left to right, because the more units we've already consumed, the less benefit we derive from each new one. Supposedly the intersection of the two curves represents the 'efficient' level of production. At any other point below this – any lower level of output – increased production increases social benefits more than social costs. At any point above it – any higher level of output – reduced production does so.

It is apparent that value considerations enter into deciding what constitute costs and benefits. Only recently have issues of long-term sustainability entered into such calculations. And there are problems in deciding what monetary values to allot to such costs and benefits. The temptation is to use the prices people are willing and able to pay for acquiring goods (and for avoiding 'bads') as a measure of the benefit of such goods (and the cost of such 'bads'). But such a 'subjective' view fails to consider (a) radical inequalities of purchasing power and (b) the gap between immediate individual assessments of welfare and actual long-term social welfare. As noted earlier, some people are willing to pay a lot for things that endanger or radically shorten their lives, for example, cigarettes.

As with traditional utilitarian calculations, a simple quantitative measure of overall social welfare in terms of the lowest cost benefit ratio also fails to consider issues of 'fair' distribution of costs and benefits. But, for the moment, we will set aside such problems and assume that we do have some sort of 'independent' measure of efficiency, in terms of which to assess the performance of 'free' markets.

Free market efficiency

The fundamental (ethical) claim of free market supporters is that the equilibrium condition of a 'free' market, and of all markets, as understood in neoclassical theory, corresponds to a socially 'efficient' situation of optimum social benefit. More specifically, the market supply curve is directly equated with a marginal social cost curve and the market demand curve with a marginal social benefit curve, and market equilibrium then equates with optimum social efficiency, the point where the marginal social cost of producing the last unit of the good is equal to the marginal social benefit from consuming it, with the best use of resources in order to satisfy human wants and needs with minimal cost.

The problem, as Hahnel points out, is that neoclassical theory itself provides good grounds for rejecting any such claim. For while market supply 'captures and represents … the costs born by the actual sellers of goods' and market demand 'the benefits enjoyed by the actual buyers of goods',[11] these 'private' costs and benefits will not necessarily correspond to social costs and benefits. For, as noted earlier, there will be individuals other than the sellers who bear part of the cost of increased production, and individuals other than the buyers who enjoy part of the benefits of increased consumption.

Hahnel uses the example of the automobile industry to illustrate the significance of such so-called '*externalities*':

> If the corporations making cars in Detroit also pollute the air in ways that cause acid rain, the costs that take the form of lost benefits to those who own, use, or enjoy forests and lakes in Eastern Canada and the United States will not be taken into account by those who make the decisions about how many cars to produce.[12]

Similarly, consumers' use of cars contributes to global warming, urban smog, noise pollution and congestion, indicating that 'the social benefit of consuming another car is less than the private benefit'.[13] Thus, the marginal social benefit curve lies somewhere below the market demand curve, and the marginal social cost curve somewhere above the market supply curve, with the two intersecting well to the left of the market equilibrium position. Such an equilibrium therefore represents production and consumption of far more cars than is socially efficient or optimal.[14]

Neo-liberals and libertarians respond by maintaining that such externalities are rare in the real world. But, as Hahnel points out, such

a claim makes little sense in face of the obvious interdependence and interconnectedness of human and natural worlds. He quotes Hunt:

> most of the millions of acts of production and consumption in which we daily engage involve externalities … Since the vast majority of productive and consumptive acts are social, that is, to some degree they involve more than one person, it follows that they will involve externalities.[15]

Hahnel notes that 'by ignoring negative external effects', markets lead to the production and consumption of more dangerous and polluting goods 'than is socially efficient'. 'By ignoring positive external effects' markets lead to less production and consumption of goods than is socially efficient (in cost benefit terms).[16]

> And while markets provide reasonable opportunities for people to express their preferences for goods and services that can be enjoyed individually with minimum transaction costs they do not provide efficient means for expressing desires for goods that are enjoyed or consumed socially or collectively like public space and pollution reduction. Markets create free-rider disincentives for those who would express their desires for public goods individually, and pose daunting transaction costs for those who attempt to form a coalition of beneficiaries.[17]

Competitive pressures will drive producers to manoeuvre 'to appropriate a greater share of goods and services produced by externalising costs and internalising benefits without compensation' in order to increase profits.[18] 'In other words, markets have an anti-social bias.'[19]

In terms of distribution of necessary costs, as well as benefits, there is no reason to expect any 'even' distribution of negative externalities. On the contrary, it will always be the poor, living down hill, down wind and down river, forced to work and live in dirty, dangerous, noisy and polluted conditions, who are the major victims of corporate cost externalisation, particularly the poor of the third world. They cannot afford to protect themselves by moving to safer areas or safer jobs, reducing stress to boost their immune systems, insulating and protecting their homes, filtering their drinking water.

Nor is there any reason to identify such externalities as in any way necessary. Effective waste treatment systems have been developed to prevent the discharge of pollutants into the natural environment, to recycle and reuse all elements of the production process, but as their installation would cut into corporate profits they have not been widely used. Again, the market fails to register the real costs of production.

Neo-liberal responses

So far we have gone along with neoclassical theory in assuming, as a first approximation, that market supply 'captures and represents … the costs born by the actual sellers of goods' and market demand 'the benefits enjoyed by the actual buyers of goods'.[20] However, as we will see, serious consideration of the real social costs of capitalist market relations requires us to significantly modify this assumption.

The most obvious response to the issue of negative externalities in a capitalist market system is substantial government (legal) intervention to try to prevent corporations from causing such damage in the first place, and ensure that they pay the costs of cleaning up afterwards. This would seem to require governments to establish relevant norms of workplace and environmental health and safety, monitoring corporations' adherence to the norms in question, with substantial penalties inflicted upon those who fail to live up to the established standards, over and above the full costs of 'clean-up'. Producers would then, presumably, pass on the increased costs of production to consumers. But, at the same time, in a competitive market, there would be pressures upon them to develop new, cheaper, cleaner, healthier, less polluting and less dangerous techniques of production.

Perhaps not surprisingly, these ideas receive little sympathy from neo-liberals and libertarians, who are hostile to any rigorous government 'interference' in private business operations. Neo-liberals are anxious to try to 'prove' that market forces alone, including free and rational negotiation between private businesses, can 'solve' the 'problem' of externalities – particularly pollution issues – without the need for any 'external' regulation.

They have a major problem, insofar as it is free market forces that have created the problems in the first place. But Chicago School economist Ronald Coase has an answer to this. He argues that the 'real' problem is not actually that of capitalist private property and capitalist market transactions at all, but rather the lack of clearly established rights of private property ownership in some areas.

The problem is that some pieces of land and some resources are not actually privately owned by individuals or corporations, with clear interests in 'protecting' them and clear legal rights to do so, including rights to recover for economic loss. With no-one thus specifically committed and empowered to fight to keep them clean, such land and resources inevitably become repositories for others 'externalised' garbage.

The solution is clear: all the resources of the world must become private property, preferably the property of powerful corporations who can make the 'best' productive use of them. In this situation, the sum of social wealth will be maximised, with minimisation of pollution costs, to produce the most 'efficient' outcome.

Coase recognises that property rights are actually bundles of different rights to exercise different forms and degrees of control over, and benefit from, resources, including excluding others' free access to such resources. But this just allows for a range of different possible ways of dividing up such resources between different owners. In fact, his basic contention is that, as far as externalities are concerned, it does not matter how such resources are apportioned, providing only that they become someone's private property.

Coase says that any private property ownership is better than none. But this allows for radically unequal distribution of property, and therefore also of the benefits of property ownership. Public ownership, on the other hand, allows for equal benefit of all, as well as democratic decision-making with respect to uses of the resources in question, and public oversight to ensure public health, safety and sustainability.

Coase illustrates his principle by reference to simple hypothetical models, rather than actual empirical data. They typically involve adjacent pieces of land used for different productive purposes by different groups of people. Some production process on land A causes land B to be polluted, thereby interfering with production (or otherwise causing harm, reducing property values, etc) on land B.

The cost of anti-pollution equipment for land A is $X; the loss to users of land B as a result of the pollution is $2X. If land B is owned by those suffering the pollution effects they can call upon the forces of the state to protect their private property. Rather than having to compensate them $2X for their lost production, the owners of land A, fully informed of the cost of such damage beforehand, will buy the anti-pollution equipment for $X. If land B is owned by the same people as own land A, they will not have to compensate the victims. But the victims (also fully informed of the costs of different options) will be willing to buy the anti-pollution equipment themselves, as the price of continuing to use the land to generate their $2X of goods. Either way, the pollution gets cleaned up, with minimum 'social' cost: the 'efficient' solution.

In reality, the mere fact of private ownership of a resource in no way guarantees that the owners will 'protect and preserve' the

resource itself, merely 'externalising' damage and pollution – into others' property and territory. On the contrary, private owners around the world continue to devastate their own property, along with that of their neighbours and others far away.

Farmers continue to pollute land and water with pesticides and fertilisers, causing soil erosion and salination through destruction of ground cover and over-irrigation. Mining companies destroy, pollute and poison their own land and water, as well as others' territories for miles around.

> Developers destroy natural habitats through yet more urban and suburban sprawl. Timber companies still plunder their great landholdings without sustainable reforestation programs … In South America, rain forests are cleared to make way for other profitable activities like farming and grazing …[21] These situations arise because the profits received from spoiling the land are highly individualised and short-term, but the costs of spoiling [it] are socialised and long term.[22]

Competitive markets, in rewarding short-term profits and punishing long-term planning, encourage an attitude of extracting the greatest wealth in the shortest term and then simply getting out and leaving others to clean up the mess. At the same time, sustainable techniques can be more costly and difficult to initiate, putting them beyond the reach of poorer, smaller business operations, without government assistance. 'Social institutions like government are therefore much better suited to prevent this type of destruction.'[23]

Following up the implications of Coase's model, we see that as the gap narrows between the cost of avoiding the pollution and the cost of damage on land B, so does the incentive for purchase of the anti-pollution equipment decline. Once the cost of the former exceeds that of the latter, then there is no longer any incentive to install it, either on the part of non-owning B-users, who lose more installing it than simply accepting their loss due to the polluters, or on the part of non-owning polluters who now pay less than the cost of the equipment in damage compensation.

Coase maintains that it would be unjust and inefficient for the polluters to have to pay more than the cost of the victims' immediate loss. They would suffer more than the victims. This is an apparent utilitarian defence of continued pollution. Unless pollution causes quantifiable damage to specific private property owners, to a value greater than the cost of clean-up, it will, and should, continue.

Coase and his followers have no problems with the fact that their

'theory' justifies continued pollution. In fact they claim to provide a quantitative assessment of what they call a 'socially-optimal level of pollution'. This is derived from a sort of supply and demand graph, similar to those considered earlier. An alleged or assumed increasing cost of each successive unit of pollution abatement by the polluters yields a rising Marginal Abatement Cost (MAC) curve, while decreasing marginal benefit of each successive unit of pollution abatement to victims yields a falling Marginal Damage Cost (MDC) curve. The net benefits of pollution abatement are maximised where MAC = MDC, at the 'socially optimal level of pollution'.

Problems of the Coase approach

There are many problems here, quite apart from this justification of continued pollution. It is misleading market values – including monetary costs that misrepresent real costs – that have, in significant part, created the problem of externalities in general and pollution in particular. The problem cannot be corrected by another set of misleading market values – in the form of the monetary value of lost production on land B. Pollution is, by no means, merely a matter of immediate financial loss through short-term disruption of production and loss of income for specific individuals on properties bordering the source of the pollution. It causes long-term harm, including genetic damage, to future generations and to the planetary environment, including depletion of finite resources, accumulation of toxins through the food chain, loss of biodiversity, global warming, ozone depletion and climatic disruption.

A true basis for beginning to calculate environmental costs is supplied by John Bellamy Foster. He defines sustainable development in terms of:

> the rate of utilisation of renewable resources … kept down to the rate of their regeneration; the rate of utilisation of non-renewable resources [not exceeding] the rate at which alternative sustainable resources are developed; and pollution and habitat destruction [not exceeding] the 'assimilative capacity of the environment'.[24]

Real costs are the costs of ensuring sustainability, and repairing the damage caused by failing to pursue sustainable – and safe – strategies earlier. Such costs could be measured in terms of use of human labour time and other resources. Coase's theorem says nothing about real costs in this sense.

Coase's libertarian emphasis upon 'free choice' holds the victim just as responsible for their loss as the polluter. They have chosen to locate their vulnerable production process (or home) next to someone else's dangerous one.

This idea has found its way into numerous microeconomics textbooks. If no property owners are directly and immediately affected, it is apparently of little significance that the world is full of dirt and noise. And the poverty that forces people to live in close proximity to others' toxic discharges is disregarded.

Similarly, Coase's 'subjective' version of utilitarianism accords equal legitimacy and value to the happiness – or monetary benefit – derived by the polluter from their environmental vandalism as to the happiness of the victim, derived from their own, responsible non-polluting activities. An objective version of utilitarianism, focusing rather upon objective measures of human health and wellbeing, is likely to produce rather different assessments of general social welfare.

Standard objections to Coase include reference to resources that cannot be divided or demarcated into private property, but are nonetheless the medium of pollution, most obviously air, and to population increase as major cause of environmental problems not obviously related to property rights.

Coasians try to address the first point by reference to individual 'rights to clean air', understood only as saleable rights to collect damages if you are forced to breath – or otherwise use – polluted air. You have lung cancer. But don't worry. You will be compensated. This only throws into sharp relief other standard objections.

The first of these centres upon radical imbalances of power and information. Major polluters include transnational corporations, with vast resources of expert legal representation and information gathering and power to bribe and blackmail government authorities and influence public opinion through media manipulations. Typical victims include millions of ordinary people, often the poorest and least informed, lacking in effective access to legal or government assistance, and to relevant information. Free and fair negotiation and agreement is meaningless in this sort of situation.

The second, and closely related, objection centres upon the 'transaction costs' of organising and negotiating such 'free' agreements. As Coase acknowledges, in situations involving lots of separate individual victims, or drawn out legal disputes between equally powerful parties, the costs of agreement could easily outweigh the benefits.

The nature of pollution is such that typical cases involve many individual victims, spread over large areas. Even where victims do have resources of money and information to effectively contest their victimisation it is easy to see how difficult and costly their effective co-operation could be, including the difficulty of achieving a united front in relation to the polluters. And even with such agreement, the task of 'proving' causation of injury can be very difficult.

As Kangas points out:

> there are thousands of pollution sources all around you, ranging from fixed factories to moving vehicles, and the weather blows their pollution around chaotically. Furthermore, different pollutants, when combined, become synergistic. So who is actually harming you? Who should pay, and by how much? This is the stuff of legal battles.[25]

Here we are in the realms of complex and expensive scientific research and scientific testimony. But, as Kangas notes, even relatively straightforward issues can be effectively obscured by clever lawyers and complex and inefficient legal procedures. 'We need look no further than the insurance industry and tobacco companies to see how even open and shut scientific cases can be disputed for decades',[26] leaving victims and their lawyers destitute and unable to continue with legal proceedings.

The problem with not monitoring – and stopping – pollution at source is that it can and does spread out around the world, interacting in complex ways with other toxins and other substances and affecting ecosystems and people in an increasingly destructive fashion into the far future. It can be carried by air and water, which cannot and should not be divided into individual 'parcels' of private property.

Say's Law

A fundamental issue of market efficiency concerns effective use of available resources of human labour power: human productive skills, strengths and abilities. This is of crucial importance insofar as people rely upon income from selling such labour power for their subsistence.

Classical and neoclassical economists have consistently argued that free markets 'naturally' tend toward 'full employment' equilibria. According to Say's Law, production creates its own demand. The money spent producing goods – wages, raw material and technology costs – provides the income with which the goods in question can be purchased by workers and suppliers, so there can be no general failure

of effective demand. Even money saved will be put into banks that will lend it for productive investment. Such banks will offer interest rates that balance supply of savings against investment demand. Those unhappy with rates that investors are prepared to pay will simply spend their money directly, rather than saving it. So each round or cycle of production generates the revenue to finance the next one. High levels of employment, once achieved, should persist.

There are major problems here, as highlighted by Marxist and Keynesian economists. First, there is the question of how full employment is achieved in the first place. Situations of less than full employment could be self-perpetuating, with inadequate demand to drive further economic expansion. To see this, we need only consider the world market today, with stagnant demand and 20–25% unemployment throughout the developing world.

Second, capitalist production is driven by pursuit of profit. The entrepreneur buys labour power, raw materials and technology and puts them to work with a view to generating more value at the end of the process than is expended in the production process itself. It is this new value, realised as profit, put back into the production process, that is the basis for capital accumulation, and the growth of the system.

Marxists highlight the fact that workers produce goods of significantly greater value than their wages, which typically cover only the costs of their subsistence. They see tools, machinery and raw materials as merely transmitting their own – embodied – labour time value (the human labour time that has gone into their own production) to the new commodities they are used to produce (in the course of their productive lifespan). But the crucial issue is the creation of new value, over and above the reproduction of the means of value creation themselves, through the interaction of people, tools, machinery and natural forces and processes of various kinds. This translates into the production of goods with greater value than the value of the goods used up in their production.

Therefore, the income generated in production will never be sufficient to pay for all the new products in an efficient and productive capitalist economy. This means that either some businesses will inevitably have to go under in order to allow others to remain profitable, or that there needs to be an expansion of the money supply in keeping with such real material expansion of the system. To the extent that such money is supplied – as credit – by the banking system, this,

again, suggests a fundamental instability in the system. Such credit temporarily fills the gap between output and income, but will have to be continuously expanded as debts are repaid with interest, with increasing likelihood of default as time goes by. In the longer term, again, some businesses – and jobs – will inevitably be sacrificed.

Quite apart from this, there will be leakages from the funds generated in production that do not find their way back into consumption – or do not do so in time to stave off a crisis of under-consumption. Money paid out in one market is spent in another, possibly far away. And some money paid out is hoarded by speculators expecting future increases in interest and bond rates.

This general principle is amply confirmed by more than 200 years of capitalist development characterised by a continuous series of cycles of booms and slumps. Investment proceeds at an increasing rate in the boom, with demand in one sector creating demand in others. The first to apply new, more productive technology are able to undercut their competitors and reap monopoly super-profits, and therefore others rush to catch up. Many, though by no means necessarily all, unemployed workers, are drawn into the work force, and into expanded consumption.

But as the new technology is generalised, super-profits disappear. And with increasing scarcity of resources in the boom, so do costs increase with profits squeezed. It becomes apparent that consumption is lagging behind production with accumulating unsold stocks. Debts are not repaid. Businesses start to collapse, others' confidence is shaken and output is curtailed. Reduced demand in one sector reduces demand in others in a negative multiplier of contraction and collapse, with increasing numbers of people thrown out of work.

Classical and neoclassical theorists have argued that the slumps are self-correcting. Declining demand for investment loans leads to falling interest rates; increasing unemployment to reduced labour costs, with raw materials and technology also reduced in price due to reduced demand and fire sales of the resources of bankrupt companies.

Even those without any income draw upon savings to maintain their consumption – and some have substantial savings – so that demand doesn't dry up completely. At some point production costs fall far enough (with possible new, cheaper, more productive technology also becoming available) to spark renewed investment and the boom is started once again. The survivors build upon and incorporate the devalued resources and markets of the fallen.

Such economists disregard the massive hardship and suffering for large numbers of working people – thrown into destitution and desperation, along the path to recovery. They ignore the possible political consequences of such suffering, as people take action to try to relieve their situation.

Equally important, as JM Keynes pointed out, is the fact that business people will not take advantage of such reduced costs unless they anticipate a substantial increase in monetarily effective demand in the near future. This is not likely with high levels of unemployment – of people and other productive resources. And there is no reason why the crisis should not simply deepen, particularly with governments cutting back their own spending – including welfare spending – to try to 'reduce the burden' on business.

There is plenty of evidence to show that periodic (business cycle) downturns are inevitable results of the operation of 'free' market forces, with unregulated overinvestment in the boom leading to overproduction, crisis and slump. Furthermore, 'free marketeers' explicitly reject the sort of Keynesian government demand management that contributed to reducing such cyclical fluctuations in the past and maintaining near full employment levels in the capitalist democracies throughout the postwar boom period.

Neo-liberal proponents of 'free markets' have embraced ideas of a 'natural rate of unemployment' necessary in order to control inflation, avoid 'overheating' and maintain economic 'health' and 'efficiency'. The basic idea is that of a residue of individuals unwilling to work at wage levels industry is able to offer, given existing technology and levels of demand. The claim is that efforts to get these people back into employment – through government spending, reduced taxation or reduced interest rates – will only produce inflation in the longer term, with no increase in employment.

Such spending will increase demand for labour, with its price increasing in consequence, thus persuading these people to enter the work force. But the declining marginal productivity of labour (with increasing numbers of workers utilising fixed amounts of fixed capital – and getting in each others way), along with increased labour costs, will mean higher prices for consumption goods, reducing the real value of the wage, and causing such people to leave the work force once again. And those still employed will now expect ongoing inflation and factor this into their wage demands, generating continuous 'cost push' inflation.

In fact, periodic downturns leave substantial productive resources unutilised or under-utilised, and there is good reason for doubting the declining marginal productivity of labour, which is a central pillar of the neoclassical theory. Modern large-scale production facilities are designed to allow for expansion of the work force up to full capacity without declining marginal productivity. And increased output from the producer goods sector can allow smaller businesses to expand their fixed capital along with their labour forces (again, counteracting declining marginal productivity). So it is quite possible for output to rapidly expand in line with increased aggregate demand (created by government spending on infrastructure), up to the point of full employment of all productive resources, including labour, without significant inflation.

It is true that once the point is reached where actual GDP equals potential – full employment – GDP, further increased demand will produce demand-pull inflation. So government economic planners need to intervene, with reduced spending, increased company taxation and interest rates at the appropriate time.

In light of these considerations we can see that the real aim of the 'natural rate' theory is to boost profits by maintaining a reserve army of unemployed people, keeping continuous pressure on the wages and conditions of those who are employed. The ever present threat of permanent job loss convinces workers to accept low wages, long hours and terrible conditions.

Of course, low wages reduce aggregate demand, investment and job creation, with increasing reliance upon debt for basic subsistence, particularly in relation to housing. So they act as a break upon economic growth, including the development and introduction of new, more productive technologies.

Not all unemployment is due to inadequate aggregate demand. It can also be 'structural' in the sense that workers lack the skills in demand at the time. But, again, this calls for substantial government action to correct the situation, through public provision of training in the skills in question.

The tendency of the rate of profit to fall

Karl Marx identified what he called a 'law of the tendency of the rate of profit to fall' underlying both shorter-term crises in profitability and longer-term stagnation and deepening crisis in the world capitalist

system. Marxists suggest that this analysis offers a more fundamental critique of alleged capitalist efficiency. They argue that the problems of under-consumption, discussed above, are the surface appearance of deeper contradictions within the capitalist mode of production itself.

We have already considered the way in which pursuit of productivity gains leads capitalist innovators to introduce new technology into the production process. Such innovators' surplus profits last only until others catch up in accessing the new technology, at which time average prices are driven back down toward the new reduced production costs.

Marx argues that the process contains the seeds of an inescapable contradiction. Marx argues that only 'living labour' – the productive life of the working class – is actually capable of generating new value, over and above the cost of its own production, whereas 'dead labour' or means of production – tools, machinery, infrastructure – merely transfers its own, embodied labour time value to the products it contributes to producing in the course of its productive life.

The cost of production of labour power – of the productive capacity of the working class – is simply the cost of subsistence and reproduction of that class, as people ready and able to work, including food, housing, practical education and transport costs. The cost of production of things, including both means of consumption and means of production, is ultimately reducible to the quantity of human labour time necessary upon average to produce them.

From Marx's perspective, the source of capitalist profit lies in forcing the work force to create new value, over and above the costs of their own subsistence and the replacement costs of the means of production and raw materials used.

The pursuit of productivity gains, followed by catch-up on the part of the rest of the industry, inevitably increases the ratio of 'dead' to 'living' labour. In the course of time, each worker comes to use more and more means of production. As Harman points out, 'for the capitalist … spending on the means and materials of production grows much faster than spending on employing workers'. So the very process of capital accumulation involves an increase in the ratio between the two, between (what Marx calls) 'constant capital' and 'variable capital'.[27]

It is the profit per unit of investment that is crucial to capitalist success, rather than the total profit achieved. And the source of profit is the surplus value produced through the exploitation of the working

population. But if the level of investment in living labour rises more slowly than total investment, the rate of profit per unit of investment will tend to fall.

As Nick Beams explains, Marx did not maintain that the rate of profit everywhere and always continues to fall. But, like all laws of capitalist production, the law of the falling rate of profit operated as a tendency that continually exerted itself. He described this tendency as the 'most important law of political economy' above all from 'an historical point of view'. This is because the tendency of the rate of profit to fall is the driving force behind the constant revolutionising of technology and other productive forces, the chief means by which capital attempts to overcome its effects.[28]

Marx acknowledged significant 'countervailing tendencies' opposing pressures towards capitalist crisis. Most obviously, the increased productivity achieved through use of new and more productive technology will continuously reduce the labour time cost of both labour power and technology itself, thereby counteracting the tendency of the rate of profit to fall. And the rate of exploitation of labour can also be increased through increasing the duration or intensity of work and cutting real wages.

Marx allows that these and other factors can indeed temporarily slow down or even reverse the falling rate of profit in the short term. But there are limits to these factors, so that the tendency will still assert itself in the longer term. Summarising this result, Marx wrote:

> The larger the surplus value of capital *before the increase of the productive force*, the larger the amount of presupposed surplus labour or surplus value of capital; or, the smaller the fractional part of the working day which forms the equivalent of the worker, which expresses necessary labour, the smaller is the increase in surplus value which capital obtains from the increase of productive force. Its surplus value rises, but in an ever smaller relation to the development of the productive force. Thus the more developed capital already is, the more surplus labour it has created, the more terribly it must develop the productive force in order to realise itself in only smaller proportion … The self-realisation of capital becomes more difficult to the extent that it has already been realised.[29]

Many economists contend that there are problems with the Marxist account. They contest Marx's conclusion that living labour is the only source of new value. And they say that Marx has trouble explaining precisely how labour time values 'determine' the actual market prices

of commodities. It is impossible to review and evaluate this debate here.

But Marx's analysis could be seen to have been borne out by the economic problems that have dogged world capitalism since the 1970s. Beams reviews the changes that followed World War II.[30] In the final analysis, the period of global expansion from 1945 to 1971 rested upon the increase in the mass of surplus value made possible by the extension of the more productive assembly line methods of mass production, first developed by American capitalism, to Europe and Japan. This brought an enormous increase in the productivity of labour and, consequently, an increase in the average rate of profit – benefiting the more efficient and less efficient firms alike – leading to further investment, expansion of industry and employment. Out of the growing mass of surplus value, capitalist governments were able to finance social welfare spending and other concessions to the working class.

However, empirical data points to the re-emergence of the tendency of the rate of profit to fall by the end of the 1960s. By 1974–75, this produced the deepest recession since the 1930s. Big business has responded to the re-emergence of falling profit rates in two interconnected ways: it has undertaken a continuous drive against the living standards and social position of the working class, and it has initiated a global reorganisation of production based on new computerised technologies.

But the development of these new methods has failed to produce an increase in the overall mass of surplus value because the previous development of labour productivity (from the steam engine to the assembly-line) has reduced necessary labour to a relatively small fraction of the working day. Even a vast development of the productive forces will only produce a tiny proportionate increase in surplus labour. Consequently, the development of new technology is unable to increase the average rate of profit as it did in the past.

In the postwar period, economic expansion brought an increase of secure, relatively well-paid jobs. Today the situation is the reverse as major corporations maintain their profits not through expansion of production, sales and employment but through downsizing and cost cutting. Every social statistic shows the same trend. A Marxist analysis suggests that these differences are rooted in the underlying crisis of surplus value accumulation, produced by the tendency of the rate of profit to decline.

Discussion topics

1. How do we measure efficiency?
2. Are free markets efficient?
3. What are externalities? Why are they a problem?
4. What are the problems with the Coarse theorem?
5. What are the problems with Say's Law?
6. Discuss the tendency of the rate of profit to fall.

Additional resources

N. Beams, *The Significance and Implications of Globalisation*, Mehring Books, Sydney, 1998.
R. Hahnel, *The ABC's of Political Economy*, Pluto Press, Sydney, 2002.
C. Harman, *Explaining The Crisis*, Bookmarks, London, 1984.

Chapter 14

REGULATING THE MARKET

Corporations and corporations law

All too often, the economic analysis of law has focused upon economic theory without reference to the reality of contemporary market relations. In particular, it has neglected consideration of the reality – if not the theory – of the corporation – as principle vehicle of capitalist production and accumulation in the modern world.

Such corporations are legal creations, through charter, prescription or legislation. Once created, typically through payment of a small registration fee, they are deemed to have the legal capacity and legal powers of natural people. They can acquire and exercise legal rights and assume legal liabilities, through contractual and other arrangements.

Such developments are supposedly justified in terms of efficiency of market operations. In particular, the limited liability of investors, liable only to the extent of the value of their investment for any obligations or penalties incurred by the corporation, making such investment more attractive than other forms of investment, is justified in terms of the greater efficiency of such corporations, compared to other forms of productive organisation.

Similarly, we must consider the limited liability of senior managers (directors, and officers and executives), deriving from the legal identification of their thoughts and acts with those of the corporation, considered as a person in its own right – the 'Identification Doctrine' – so that they have no (or very limited) personal responsibility for the consequences of such thoughts and actions in the wider community, being

personally answerable only to their own shareholders. Such limited liability on the part of senior managers is justified in terms of its encouragement of the 'most talented' people to apply such talents to effective corporate direction. If senior managers were personally responsible for their corporate decision-making (to anyone other than their own shareholders) clever people would refuse to take up such positions. Perhaps no-one would do so, and we would all lose out in the resulting reduction in economic efficiency.

Because investors can limit their risk, they can invest more widely than they would otherwise have done. This allows for the formation of more and bigger corporations. The larger corporations become, the greater their productive efficiency – in terms of reduced costs – achieved through economies of scale, research and development (in the longer term), and reduction of transaction costs through 'internalisation' of necessary operations.[1]

Several problems arise. In the first place, it is not true that shareholders provide any significant proportion of the capital available to big corporations to fund productive investment. On the contrary, dividend payouts to shareholders represent a massive drain on corporate funds, far greater than the input of investment capital from new share offerings.

As Marjorie Kelly points out:

> invested dollars reach corporations only when new equity is sold. In 1999 the value of new common stock sold was $106 billion, whereas the value of all shares traded was a mammoth $20.4 trillion. So of all the stock flying around Wall Street, less than 1% reached companies … from 1998 to 1999 … the value of stocks increased $1.1 trillion, while sales of new common stock were $83 billion, or about 7% of the increase. Thus we might conclude that the market is 7% productive … Yet this leaves out … stock buybacks … For 1998 [this] figure was a negative $267 billion. Thus equity issues were ultimately a negative source of funding for corporations … And that's not even including dividends, which in 1998 extracted an additional $238 billion from corporations … New equity sales were a negative source of funding in 15 out of the 20 years from 1981 to 2000. In other words … [there is no] stockholder money going in – it's all going out. The net outflow since 1981 for new equity issues was a negative $540 billion … [the bull stock market over this period was actually created by] companies pumping massive amounts of money into it to prop it up.[2]

And Kavaljit Singh points out that, 'even new issues may not contribute to additional investment if the proceeds are used to retire other

domestic debt or fund current expenditures or are mobilised for speculative purposes by the corporation'. In fact, institutional investors in securities should much 'more accurately be referred to as institutional traders'.[3]

In the second place, it is far from clear that large size is necessarily associated with productive efficiency in the – genuinely useful – Adam Smith sense, of reducing real production costs in terms of human labour time, or conserving scarce resources. And this applies in both the technical sense – that developments in science and technology (allowing for more efficient use of small scale 'post-Fordist' production techniques) undercut older arguments about 'inevitable' economies of scale – and in the social sense that, in practice, large scale creates monopoly and oligopoly that insulate corporations from the need for productive efficiency, or denies consumers the benefits of such efficiency.

In the third place, it is all too easy for corporate executives to utilise the legal transfer of their own thoughts and actions to the corporation to engage in anti-social actions without fear of effective penalties.

As Glasbeek points out, this is particularly clear in the case of small corporations, sometimes created purely for the purposes of individual evasion of taxes and bills, compensation payouts and workers' entitlements, with corporate bankruptcy – and covert transfer of assets to other corporations – utilised to evade such payments.[4] But the consequences of such irresponsibility on the part of senior managers of big corporations can be of global social significance, destroying lives and environments on a huge scale.

All around the world, corporations law centres upon the legal duties of managers and directors to make money for shareholders, as owners of the corporation. Shareholders can sue such executives for failing in this duty. Such law fails to consider any idea of broader public interests and responsibilities associated with the huge powers exercised by such executives. Thus, as Bakan says, 'corporate social responsibility is illegal'. Here, indeed, competitive 'market forces' (which reward short-term profit maximisation and severely punish anything else) and corporate law work together to 'inhibit executives and corporations from being socially responsible'.[5]

In the fourth place, the effective separation of ownership and control within the corporation offers huge opportunities for directors and managers to pursue their own interests at the expense of the shareholders themselves (as well as the rest of the world). This

includes a spectrum of possibilities, from short-term profit maximisation at the expense of future profitability with a view to cashing in their own shares quickly and moving on, to complex, Enron-style, 'deceptive schemes to take the investing public to the cleaners' with auditors acting as 'the deceptive managers lapdogs'.[6]

Law and economics theorists argue that there is no loss of effective control by shareholders. For while corporate law specifies directors' duties to shareholders only in broad and general terms, the formation of a corporation involves an implicit contract between investors and managers, which could, in theory, be made explicit and could specify greater control by investors.

In fact, as Glasbeek says:

> directors and managers, charged [only] with pursuing 'the best interests of the corporation', find it legally easy to justify almost any choice they make on how to use the capital put into their hands. Indeed, courts are at pains to say that they will try to respect their judgment, provided that it appears to have been exercised honestly. Now, if human beings are believed to be essentially egoistic [as in both classical and neoclassical market models] it is to be expected that, with the wealth of choices available, many decisions will be made that make it possible for directors and managers to shirk work, while giving them more prestige and pay – even if this comes at the expense of investors/owners.[7]

Free market liberals argue that the market itself will 'discipline' managers, insofar as they are rewarded (or punished) in proportion to changing share values. But as Glasbeek says:

> in practice, these potential market incentives and disincentives do not work very well. For one thing, it is difficult for investors/owners to know what a non-abused corporation's best results would be and, therefore, how to discount the value of shares accurately for past and future bad managerial behaviour. For another, investors/owners cannot easily judge the extent to which the decision-making by the directors and managers alters a corporation's performance. This problem reduces to guesswork any market discipline to be exercised by changes in share valuation.[8]

In practice, directors and CEOs seem to be in a position to judge their own performance and determine their own rewards. They can manipulate the balance sheets of their companies, and thereby manipulate share prices, allowing them to choose when to cash in their share options and at what values.

Real-world cost benefit analysis

In the real world of capitalist market relations – as opposed to the models of law and economics theorists – corporations (or rather directors and managers, sheltering behind the 'corporate veil') perform cost benefit analyses all the time. As Joel Bakan points out, in his book *The Corporation: The Pathological Pursuit of Profit and Power*, such analyses 'are at the heart of corporate decision-making'.[9] Far from utilitarian concerns for 'the common good', the issue is the financial good of the corporation itself, understood first and foremost as the interest of corporate shareholders (or of the managers).

> The corporation's institutional makeup, its [legal and practical] compulsion to serve its own financial interests above everything else, requires executives to make only those decisions that create greater benefits than costs for their corporations.[10]

Harm to others, namely, 'workers, consumers, communities, the environment',[11] figure in the equation in terms of expected monetary costs incurred through application of existing criminal and civil law, penalties for breaches of health and safety or pollution control regulations, and compensation payouts to victims resulting from civil litigation. This includes the cost accorded by the legal system to such externalities as lost human lives, horrific injuries and significant environmental degradation.

Over the years, such cost benefit analyses have had terrible consequences. Bakan considers the General Motors (GM) fuel tanks case in 1993, involving horrific burns to a woman – Patricia Anderson – and her children due to dangerous placement of the fuel tank in her Chevrolet Malibu.

> The evidence in the trial showed that General Motors had been aware of the possibility of fuel-fed fires when it had designed the Malibu and some … other models. Six fuel-fed fire suits had been filed against the company in the 1960s, 25 more in the early 70s, and in May 1972, a GM analyst predicted that there would be another 60 by the mid-70s. On June 6, 1973, around the time GM began planning the … Malibu that Patricia Anderson was driving, GM management asked a [company] engineer to analyse [such] fires in GM vehicles … In his report, [he] multiplied the 500 fuel-fed fire fatalities that occurred each year in GM vehicles by $200 000, his estimate of the cost to GM in legal damages for each potential fatality, and then divided that figure by 41 million, the number of GM vehicles … on US highways. He concluded that each [such] fatality cost GM $2.40 per automobile … The

> cost of [safer] fuel tanks [$8.59 per car] meant that the company could save $6.19 per [car] if it allowed people to die … rather than alter the design … [12]

The Los Angeles Supreme Court found that GM had dangerously positioned the fuel tanks to save costs and the jury awarded compensatory damages of $107 million and punitive damages of $4.8 billion. GM appealed to the California Court of Appeals, where the US Chamber of Commerce – 'a leading voice of big business'[13] – described the previous jury decision as 'illegitimate' and argued that the kind of cost benefit analysis applied by GM was 'the hallmark of corporate good behaviour'.[14]

Such legal payouts appear to have failed to deter ongoing corporate wrongdoing – and social harm – on a huge scale. According to Bakan, 'corporate illegalities are rife throughout the economy. Many major corporations engage in unlawful behaviour, and some are habitual offenders with records that would be the envy of even the most prolific human criminals.'[15] GE, the world's largest corporation, was found guilty of 42 major legal breaches between 1990 and 2001, including repeated contamination of soil and water with highly toxic PCBs, defrauding government on defence contracts, violating worker safety rules, illegal sale of fighter jets to Israel, deceptive advertising, overcharges on mortgage insurance, safety violations at nuclear fuel plants and failure to properly test aircraft parts. Some of the fines were in the tens of millions of dollars; one was $2 billion.[16]

Bakan concludes:

> the corporation's unique structure is largely to blame for the fact that illegalities are endemic in the corporate world. By design, the corporate form generally protects the human beings who own and run corporations from legal liability, leaving the corporation, a 'person' with a psychopathic contempt for legal constraints, the main target of criminal prosecution. Shareholders cannot be held liable for the crimes committed by corporations because of limited liability … Directors are traditionally protected by the fact that they have no direct involvement with decisions that may lead to a corporation's committing a crime. Executives are protected by the law's unwillingness to find them liable for their companies' illegal actions unless they can be proven to have been 'directing minds' behind those actions. Such proof is difficult … to produce in most cases, because corporate decisions normally result from numerous and diffuse individual inputs …[17]

Even where directors have been found to have been the 'directing minds' for purposes of civil liability, this has typically rendered the

corporation liable, rather than the directors themselves. And 'losses' in the form of compensation payouts have been readily passed on to insurers, the tax office and consumers. Furthermore, in recent times, in the area of the law of negligence, there has been an increasing push to restrict corporate liability, abolish punitive damages and radically limit payments to victims of injury and illness. We will return to this point in the chapter on torts.

Monopoly and oligopoly

While 'free' markets might have major externality problems, they appear to have obvious advantages for at least some consumers, able to choose between suppliers competing on price and quality, as compared to situations of monopoly and price fixing (collusive tendering, etc). But ever since Adam Smith, economists have acknowledged that insofar as competition conflicts with the prime goal of profit maximisation, 'free competition' will be difficult to achieve in practice.

Businesses will inevitably strive to neutralise competitive price reductions and costly development of new technology through the formation of *cartels and monopolies*. Moreover, the process of competition will tend to be self-undermining as weaker, higher-cost competitors are forced out of business, with their markets and productive resources taken over by other, stronger firms. Periodic crises of overproduction, themselves products of unplanned market competition as all rush to cash in on new technologies and limited markets, accelerate the process of concentration as profits are squeezed and smaller, weaker competitors fall by the wayside.

The long-term tendency of capitalist markets is toward an increase in size of individual businesses, what Marx called *concentration of capital*, seeking to increase the scale of their production and thereby gain economies of scale, and a *centralisation of capital* through mergers and takeovers. And after more than 200 years, the commanding heights of the world economy are dominated by transnational mega-corporations.

Pro-marketeers argue that the survivors are those firms best able to provide consumers with products giving the optimum combination of quality and price. The producers who are most efficient, in terms of increased productivity and reduced costs, are those who win out in the competitive struggle for markets, through offering cheaper and better products to consumers. And the resulting increase in size allows for

further efficiency gains through economies of scale, increased research funding, reduced transaction costs, etc.

In fact, as Shutt points out:

> victory all too often goes to those with the deepest pockets who are best placed to withstand a competitive price squeeze. Moreover, such financial power can also be, and frequently is, exercised to prevent new competitors entering the market, whether by means of predatory pricing to prevent new entrants from gaining market share, pressurising distributors not to handle competitors' products or simply by buying them out through acquisition of their shares. There is an in-built tendency in any given product market for competition to be reduced in the long run as corporations seek to protect their investors' return on capital, to the disadvantage of the very consumers who are supposed to be the main beneficiaries of the process.[18]

This is an area where neoclassical theorists and neo-liberals reluctantly acknowledge an important role for law. And, indeed, around the western world, it is predominantly neoclassical ideas that inform, direct and legitimate legislation addressing issues of monopoly, price fixing and other 'anti-competitive and restrictive trade practices', so-called *competition policy*.

Most free marketeers argue that the market gives consumers 'freedom of choice' and competition ensures a genuine choice of high-quality, low-cost products. But profit maximisation requires low production cost and high market price, and increasing monopolisation along with inadequate consumer access to reliable information about the goods offered for sale leaves customers, particularly those with limited spending power, vulnerable to marketing of shoddy, dangerous, overpriced goods. Some libertarian free marketeers also acknowledge a place for consumer protection legislation.

In Australia, the Australian Competition and Consumer Commission (ACCC) is supposed to play a central role in both of these areas, enforcing anti-monopoly and consumer protection legislation. Part IV of the Trade Practices Act 1974 (Cth) covers anti-competitive practices that limit or stop competition. It aims to

> foster the competitive environment necessary to give consumers a choice in price, quality and service. It prohibits commercial conduct that substantially lessens competition in a market, as a lack of competition might allow some traders to push prices up and lower the quality of the goods and services they offer to consumers. Some anti-competitive conduct is prohibited outright (for example, price fixing), while

> other types are prohibited only if they substantially lessen competition. A substantial lessening of competition is apparent when the ability of buyers to shop around for a deal that suits them is significantly diminished.
>
> There are some circumstances in which a refusal to supply is unlawful under the Act. These include a misuse of market power, third line forcing, boycotts, resale price maintenance, and placing limitations on resellers.[19]

Section 51 AC deals with 'unconscionable conduct', including bullying of smaller firms by larger ones. Other sections deal with consumer protection, including misleading or deceptive advertising, failure to provide necessary safety information with products, and sale of dangerous products.

The ACCC polices all these areas. On occasions, it successfully prosecutes major price fixing or other conspiracies. A recent case involved the use of whistleblower evidence to mount prosecutions of power transformer and distribution transformer cartels, leading to record $35 million penalties against companies and senior executives.[20]

But such legal intervention has limitations, particularly in a relatively weak and dependent economy like Australia. When a handful of transnational corporations control 50 per cent or more of world production and distribution in a range of important industries (oil production and distribution, steel and aluminium production, car production, computer manufacture, food processing, laundry detergent manufacture, aircraft production and airline travel), they have great power to influence government policy.

Governments are dependent upon such corporations to provide jobs and tax revenue. Members of government – or their associates – can be shareholders in such corporations and take up positions on their managing boards upon retiring from politics.

Even if the will is there, on the part of those formulating and enforcing legislation, it is seldom supported by anything like the resources necessary for effective action. There are substantial problems in acquiring legally acceptable 'evidence' of price fixing and misinformation and in mounting an effective prosecution when challenged at every step by the political, legal and media power of big corporations.[21]

In practice, with the top 200 multinationals producing one-third of total world output of commodities and conducting two-thirds of inter-

national 'trade',[22] markets are far from the sort of freely competitive markets of neoclassical theory. Profits in more freely competitive – typically local and smaller scale – sections are squeezed by the monopoly pricing of the big international corporations. And capital market deregulation around the world opens the way for takeover of all profitable local businesses – or effective competitors – by expanding multinational corporations.

On a world scale, the norm is what is called *oligopoly*. As Stilwell says:

> This is a market structure in which there are relatively few firms in the industry, and each usually produces differentiated [brand-named] goods or services, the prices of which are directly under the control of the firms themselves. New firms have difficulty coming into the industry, either because of the inherent difficulties of establishing sufficiently large productive capacity or because the collusive practices of existing firms impose substantial 'barriers to entry' … Price rigidity … is a common consequence of interdependent behaviour … even without explicit collusion in price fixing. Prices tend to … remain fixed for quite long periods … despite fluctuations in demand and supply conditions … Periodic bouts of vigorous price competition may occur, especially when a new entrant threatens to invade the market territories of existing firms. Typically … non-price competition prevails. Rivalry over market shares, focused upon … advertising … may be intense, but … collusion is a strong tendency … firms may agree to a territorial division of the market … [or] a system of tacit price leadership, whereby one firm sets the standard and the others follow, is commonly adopted.[23]

Neoclassical theorists themselves bemoan a situation where 'a small minority of giant firms dominate the marketplace such that small producers feel relatively powerless',[24] where oligopolists 'generally prefer to collude than compete on prices, keep prices high and stable as long as conditions allow, drive out or vigorously resist entry of new competitors' and 'manage the media, public opinion and government contacts to resist any restriction on their operations'.[25]

Conversely, under certain conditions, a free competitive market is not actually the most efficient, even in neoclassical terms. Shutt explains:

> This applies in particular to those sectors of the economy requiring substantial investment in infrastructure networks – such as trunk roads, railways, water supplies, energy distribution and telecommunications. For in such cases the creation of competing networks has generally been shown to be uneconomic because of the need to undertake

> double or treble the amount of capital expenditure to provide a service which can quite easily be supplied from a single network – a consideration which until the 1980s was reflected in the general acceptance that such public services were *natural monopolies*. Likewise in the case of industries, such as manufacture of defence equipment, serving a very limited market (often comprising only their own national governments) but requiring massive investment in research and development as well as fixed capital, it is clearly highly inefficient to require competitive bids.[26]

In an economic downturn, with profits squeezed, private corporations must sack workers to cut costs, leading to reduced consumption and deepening crisis. State-run operations do not need to make profits and can reduce the severity of the crisis by continuing to maintain – or increase – jobs, salaries and consumption.

As long as such monopolies are government controlled, cost saving can be passed on to the public with limited monopoly profits available to subsidise public services. But with the neo-liberal and libertarian demand for privatisation in the name of 'market choice' and 'efficiency', such monopolies are handed over to direct increasing monopoly profits to a minority of super-wealthy shareholders and executives (with little effort to develop future services), or attempts are made to introduce 'competition' through wasteful duplication of infrastructure that similarly pushes up costs to the public.

In the past, company use of such resources as telephone lines has subsidised public access. This created an incentive for major private users to take control of such resources and force the public to subsidise corporate use of them. As Stilwell points out, application of 'competition policy' to public utilities has led to the 'relinquishment of community service obligations that previously guided the behaviour of the public enterprises'.[27] Dividends and executive salaries are maximised at the expense of updating infrastructure, leading to shortages and breakdowns such as have been experienced since the 1990s in various places, including California and New Zealand.

Whatever steps are taken in the short term to 'open up' competition, privatisation puts these resources into an open world market where they become prey to takeover by multinational oligopolies. As Beder points out, foreign owners can then cut off crucial parts of the economic system of the countries in question for political reasons.

> in 1998 when Quebec was experiencing an electricity crisis … a private US company shut down its plant until it could get the price it wanted

> for its electricity … US companies also shut down supply in the Dominican Republic to force the government to pay its debt to them.[28]

As Shutt observes:

> It is ironic that a policy which was ostensibly intended to reinvigorate the prevailing state-owned monopolies by subjecting them to the pressures of the competitive marketplace has only served to emphasise the difficulties of securing competition in a way that benefits consumers.[29]

We need to be cynical about the real reasons for such privatisation, considering the incredible political and economic power associated with control of such crucial elements of infrastructure, and especially considering the inevitable pressure upon state provision of public health and education with the removal of direct subsidies from such state monopolies. This is all the more important given the ongoing pressure for further reductions of company taxation in a world of internationally mobile big capital.

Discussion topics

1. What are corporations? Critically analyse their advantages and disadvantages.
2. What are the principal goals of corporations law? What are the problems with this?
3. Are national and international markets really 'free' markets? Include reference to oligopoly and monopoly.
4. What are the problems in achieving effective 'competition policy'?

Additional resources

H. Glasbeek, *Wealth by Stealth*, Between the Lines, Toronto, 2002.
S. Mann, *Economics, Business Ethics and Law*, Lawbook Co, Sydney, 2003, chs 5 and 11.
K. Singh, *Taming Global Financial Flows*, Zed Books, London, 2000.

Chapter 15

IS THE MARKET JUST?

Whereas the major concern of the previous chapters was the *efficiency* of 'free' markets, we now examine the *justice and fairness* of the capitalist market system. More specifically, this chapter and the next focus upon some ambiguities, confusions and problems associated with ideas of equality of reward as a basic principle of distributive justice. Insofar as such an idea has served to legitimate and guide moves away from policies informed by social-liberal ideas and values to policies driven by a neo-liberal agenda of deregulation, privatisation and free market forces, these issues take on profound practical significance.

This chapter focuses upon reward for direct involvement in the provision of goods and services – in the form of wages or salaries. Chapter 16 moves on to the neo-liberal defence of inheritance and return to productive use of privately owned assets.

Just deserts

A key concept of neo-liberal ethical, legal, political and economic thinking, is that of the 'free market' as a fair and just system of distribution of resources.[1] In this view, everyone obtains their fair share of the total social product, in proportion to their productive contribution. The neo-liberal assumption is that the free market generally ensures equality of reward or just deserts, in the sense that individuals get out of the market system what they freely choose to put into it; or appropriate and fair recompense for goods or services they offer in the market and they are free to reject anything less.

But there is a crucial ambiguity in the idea of just deserts in this context. It could mean that individuals are legitimately rewarded in proportion to the effort, commitment and sacrifice they put into socially useful production. Alternatively, it could mean that producers are legitimately rewarded in proportion to the actual value of the goods produced or services offered. The more good things or the better things (better made, more valuable things) an individual produces, the greater the reward they deserve.

As we have seen, according to the neoclassical economic theory favoured by neo-liberals, the market prices of goods are determined by the interaction of demand and supply curves. It is assumed that such interaction determines both the market price and the quantity produced, with the market settling into a 'general long-run equilibrium', where the size of each industry has adjusted to the level of demand so that prices correspond to 'marginal costs'. The market settles at the price and quantity where market supply and demand curves cross. Here, the number of units buyers want to buy is equal to the number of units sellers want to sell, and all available resources are fully and efficiently utilised.

Changes in demand due to factors other than the price of the good (for example, increases in disposable incomes) shift the demand curve up or down the supply curve. Increased demand, expressed in consumers' willingness to pay more for the good in question, will lead to increased revenues for producers and consequently increased output of the good as investment shifts to the higher-profit sector. Such a flow of investment leads to increasing competitive price reductions, with prices ultimately pushed back to marginal costs in a new equilibrium.

Fundamentally, neo-liberals support the idea of reward in proportion to the value of goods produced, as expressed in market price. Since they see the market as the only legitimate source of income (apart from gifts and self-sufficiency), it seems to follow that such market forces will limit the funds available to reward producers. And if market price reveals the real social value of goods produced, it seems reasonable to suggest that individuals should be rewarded in proportion to the amount of real social value they contribute. In particular, reward in proportion to output will encourage optimum output, with society as a whole benefiting.

At the same time, neo-liberals argue that greater effort and sacrifice will generally lead to greater output of value and so reward for output will also be reward for effort and sacrifice. Also, fewer people

will be willing to work in unpleasant, dangerous, boring and demanding jobs, and the resulting limited supply of labour will automatically lead to higher wages in these areas.

Those who are willing to undergo prolonged training and education to acquire special skills will thereby become more productive of social value, and therefore deserving of higher wages. The lower wages and greater sacrifices associated with such prolonged training will mean a limited supply of such skilled labour, again, leading to higher wages through the natural operation of market forces.

On the other hand, there is no point in individuals expending effort and sacrifice in producing things that are not really wanted by consumers. It is therefore quite legitimate that reduced demand, leading to reduced prices, should also lead to reduced wages. Such wage reductions function to tell workers that it is time for them to move out of declining industries and into growth areas producing goods of real social value. Concomitantly, it is quite legitimate to reward workers in such growth areas at higher levels in order to encourage others to move into these areas and increase output of socially valued commodities.[2]

Supposedly, everyone is 'free' to refuse any wage offer they regard as unfair or inadequate. In reality, without independent wealth, or better offers, everyone is forced to take what is offered. With 'oversupply' of labour employers can pay less than subsistence wages. Only restricted supply forces them to offer wages that reflect labour's real value-creating power. Yet the reality of the current world is massive oversupply of labour in all regions. As Shutt observes:

> a distinguished British economist estimated in 1997 that the true level of male unemployment [as opposed to government figures] was broadly similar as between the US, Britain and continental Europe – at 11% or more. At the same time] the surplus labour capacity [of the third world] would equate to an unemployment ratio of not less than 20–25% expressed on a similar basis to that applied in OECD countries, [with the position] continuing to deteriorate.[3]

In Australia, as elsewhere, the real employment situation is disguised by misleading official figures. In this context, Watson et al observe that such official figures suggested unemployment of 620 000 persons or (6.1%) in January 2003. But

> the definition of 'being unemployed' used by the ABS is very restrictive. If a person does more than one hour's work in the survey reference week, they are deemed to be employed. In addition a large number of unemployed people withdraw from the labour market.

> These discouraged jobseekers are not counted in the unemployment rate.[4]

These figures also ignore the increasing levels of underemployed workers, estimated by the ABS at 563 000 people in 2001. These are people forced to work for less hours than they want to (between 11 and 17 hours per week less), and typically working in low paid, insecure casual and part-time employment.

> The group with the largest shortfall in hours are those under-employed workers who took up part-time work because they could not find full-time work. The average hours of work sought by this group was 37 hours a week, while the actual hours they worked was just 20 hours.[5]

Adding official figures for underemployment to those for unemployment for the same period gives 1 240 300 people or 12.6% of the labour force. This does not include the hidden unemployed, and the situation has worsened since 2001.

This does not mean that these vast wasted resources could not or should not be employed to fulfil desperate need on a huge scale. Rather, it means that the profit-driven capitalist market system has failed to generate monetarily effective demand on the part of those in need to motivate the relevant investment.

Neoclassical theory maintains that oversupply of labour is an indication that such labour is overpriced, and that falling wages can restore full employment. But with wages below subsistence for vast numbers of workers, such a prescription is clearly wicked and stupid. And in a situation of inadequate aggregate demand, further falls in wages will only reduce such demand, and consequently also productive investment, still further.

It is apparent that in the real world of existing capitalist markets there is often an inverse relationship between effort and financial reward. Desperate millions struggle for long hours of strenuous, difficult and unrewarding effort in dirty, dangerous third world sweatshops for a pittance, while a non-executive director of a big first world corporation attends a couple of board meetings in a year in exchange for a hefty pay packet, and tax-free expenses that would feed thousands of such poor third world-ers.[6]

In the richest country, the United States, many of the more than 30% of the population employed in the service sector are kept on their feet all day, or night, running around until they are exhausted, for $6–7 per hour, or $2.40 plus tips for many waitresses: only 91% of what

such jobs earned 30 years ago and not enough for minimum rents of US$500–700 per month and food, let alone private medical costs, kids school fees, etc. Many have to work two jobs or more for 70 or more hours per week, just to pay for the most basic of subsistence.[7]

Confronted with these facts, neo-liberals argue that sweated labour in the third and first worlds is of little value, due to the low value of its products. Such workers are unskilled and therefore comparatively unproductive, contributing little of real social value. Their output is oversupplied, and their low remuneration is therefore justified (in utilitarian terms) in communicating to them the need to produce something else – in greater social demand.

The problem is that such workers are physically prevented from doing anything else in many cases, through barriers to international labour mobility and/or very limited job and (affordable) training opportunities in their home territories. In neo-liberal terms, they are not free to choose to do anything else, no matter how much effort they make. And, as noted above, there are no jobs for third world-ers in the first world anyway, even if they were allowed access to it.

Utilitarian defence of 'incentives'

The radical disparities of reward in the real world raise other issues. Even if actual wage and salary structures do reflect significant differences in the value of output from different individuals and groups, do the welfare benefits of such disparities outweigh the welfare costs of the resulting inequalities?

The benefits allegedly arise from the 'incentive' effects of such proportional recompense, motivating continued high levels of output from which all will benefit though 'trickle down'. But, in fact, there is no solid evidence to support the idea that proportional recompense is necessary to motivate high levels of output. On the contrary, as Self points out, 'not so long ago top earners in Sweden paid over 90% in taxes without any discernable adverse effects upon the good economic and export performance of Sweden at the time'.[8] Nor is there any evidence to support the trickle down argument. On the contrary, as noted in chapter 13, substantial increases in income inequality in Britain and the US from 1980 to 1990 went along with increases in absolute poverty.

At the same time, the work of Wilkinson, also referred to earlier, provides strong support for the idea that even with declining absolute

poverty, increasing income disparity leads to increasing morbidity and mortality, drugtaking, violent crime and general loss of social cohesion, trust and stability.

It is far from clear why producers' incomes should be dependent upon changing conditions of demand for their products, quite independent of the effort, commitment and sacrifice they put into production. Why should some gain substantial wage increases just because, for example, an increase in disposable income of some other group has increased demand for their products? Why should others lose income just because of some change in public tastes, perceptions or priorities?

The typical neo-liberal response is to appeal, once again, to utilitarian considerations, saying this is the only way to balance supply with changing conditions of demand. Indeed, this is frequently held up as the central benefit of market relations. But this is not the only way to balance supply and demand. Systems of democratic economic planning provide viable alternatives, with many other advantages.[9]

In fact, workers are more likely to lose their jobs altogether as a result of declining demand for their products than to suffer reduced incomes. Nor are they likely to see any increases in income as a result of increased demand. And neo-liberals have hardly led the way in developing programmes of social intervention to reskill and reemploy those who lose their jobs through declining demand.

Perhaps most important is the issue of the real value of the output of the low-paid workers referred to earlier. Are their contributions of coffee and tea and bananas, of zinc and copper and coal, clean rooms and streets and cooked food, really so valueless compared to the generally less tangible or obvious contributions of managerial or legal expertise of those at the other end of the income spectrum?

The real value of output

We have seen that neo-liberals, strongly influenced by neoclassical economics, generally defend the idea of market prices as a true measure of the value of goods produced. Neo-liberals reject objective (scientific) assessments of the value or usefulness of products and services in terms of basic need satisfaction, the maintenance of physical and mental health of current and future generations, individual empowerment and the long-term sustainability of production practices. Rather, they favour a subjective approach, centred upon individ-

uals' 'free' choices of what they really want. No-one has the right to deprive others of the freedom to make their own decisions about what they want and what is good for them. No-one should presume to dictate to others what is really good for them.[10]

There is some truth to the idea that people, at least sometimes, know what they want and should be allowed to act upon that knowledge, as long as this does not restrict others from doing similarly. But there are limits to how far this idea can be pushed. There is plenty of evidence that people can be, and frequently are, radically mistaken in their assessments of what is good for them. And they can and do live to regret their mistakes in this area.

A person's idea of what is good for them or what they want depends, in part, upon their state of knowledge. People do not have perfect knowledge about all the goods and services available in a capitalist market system. Indeed, huge industries strive to influence their choices without any reference to real knowledge of the products in question their conditions of production, their long-term health effects, the conditions of their disposal, etc.

Further problems arise when we consider demand curves as accurate expressions of individual choices, preference and priorities. Most obviously, some might have great desires or needs for products but not the money to buy them, or buy enough of them. On the other hand, others with lots of money might generate substantial (social) demand for things they hardly care about. Demand curves depend upon available income, which is unevenly distributed in contemporary society. As we have noted, a person who desperately needs some expensive commodity (for example, an AIDS drug) is by no means necessarily in a position to earn the money to acquire it, no matter how much effort they make. People might, and do, buy lots of things only to find that they do not really want them after all. People might, and do, strongly desire things that the market does not or cannot offer, things like genuine empowerment, respect and equality.

We can see how individuals, groups or nations could be, and are, bullied or misled into entering into market exchanges that actually leave them worse off, including lost opportunities for acquiring better things.

On the other side of the coin, we have good reason for rejecting the idea of market prices as expressions of the real social cost of production of goods. Most obviously, there is the issue of monopoly and oligopoly pricing, allowing producers to push the market prices of their

goods far above their costs of production. This includes clever media brainwashing, convincing consumers that particular branded products are imbued with magical powers of various kinds that differentiate them from identical products from other manufacturers, as well as the operation of price fixing cartels.

There are numerous examples of purchasers paying less than the real social costs of production, distribution and consumption of particular products. As noted earlier, they buy cheap labour in the third world or the first world service sector for less than the subsistence costs of the work force. They purchase beef and wheat and coal from overseas suppliers for a mark-up on the monetary cost of the production of these commodities (fertilisers, insecticides, harvesting, etc). But they fail to pay for the loss of topsoil, desertification, salination, poisoning of rivers and ground water, land subsidence, ecological destruction and destruction of agricultural land through dumping of tailings that are the real costs of such production. They fail to pay for the development of alternative sustainable resources to replace non-renewable resources consumed at a rate faster than the rate of their regeneration. And they fail to pay for the health, pollution and climate effects suffered by others as a result of their own consumption.[11]

Market prices at equilibrium are supposedly already, in part, determined by labour costs, and it therefore begs the question to justify or explain such wage levels in terms of market prices. This is (partly) why the products of third world and first world service sector labour are so cheap: because such labour is cheap. The crucial question here is why it is possible for employers to get away with paying such low wages.

Non-free markets

Neo-liberals, in taking market price as a true measure of value of goods and services, legitimate the ongoing super-profits of those who contribute to maintaining situations of over or under supply, in order to keep prices and profits up in some areas and down in others. This problem is exacerbated by the failure of neoclassical theorists to provide any clear (empirical) criteria for establishing whether or not a market actually is in 'equilibrium'.[12]

Neo-liberals will condemn price fixing and monopoly as incompatible with the free market relations they support. They will condemn all barriers to the free flow of productive resources in response to

demand. But they are far from consistent on this point, as we have seen in relation to the international mobility of labour. Neo-liberals support the liberation of individual initiative in the service of economic efficiency through the free international mobility of capital, in pursuit of maximum profit. Yet, they are frequently less enthusiastic about equally free mobility of labour in pursuit of higher wages. In the world today national boundaries provide barriers to the free flow of labour in response to demand. Millions of workers are trapped in the third world in a situation of massive oversupply of labour due to ongoing displacement of farmers and others from the land by agricultural revolution, the spread of agribusinesses, big dam projects and other developments, with no chance of moving to areas of higher demand, no matter how much effort and sacrifice they make. Many who do make real efforts and sacrifices to try to improve their lot are rewarded by imprisonment and deportation.[13]

Millions of third world producers are trapped in regions where, because of a legacy of domination by foreign colonial and imperial powers, destroying indigenous industries in favour of primary production of monocultures and mining (including the directives of the International Monetary Fund, urging numerous poor countries to concentrate upon the same few export crops[14]), they have no choice but to grow crops that are oversupplied on the world market. They are locked together in a desperate competition for foreign currency that forces their wages and revenues ever lower.[15]

The products they produce are relatively cheap in part because of the barriers to free flow of labour out of the third world, along with political repression, lack of indigenous industrial development, and competition with heavily subsidised first world primary production. There is therefore no way the cheapness of their products can be used to justify the low levels of wages, as neo-liberals suggest. Meanwhile, powerful first world monopolies regulate and control the supply of high technology so as to maximise their profits on a world scale. This means that technology produced in the first world is in short supply globally compared with labour, and in high demand, including high demand from the poor third world-ers, who are competing among themselves for the foreign currency to pay for it. So the terms of international trade inevitably turn against the latter.[16]

In the first world, the poverty of those in the service sector makes it very difficult for them to seek better jobs, even in conditions of increasing demand for labour. As Toynbee says, in the United States:

> Those living hand to mouth cannot risk losing a week's pay between jobs, for fear of eviction from their rooms or hunger. Shopping around for better jobs means lengthy applications, references, humiliating drug tests and waiting – all of it taking time they cannot afford. Employers conspire to keep wages down by not advertising their rates of pay, so that no-one knows what might be available elsewhere – and there are no unions, of course. In cities with little public transport, people rely on cycling or lifts from friends. They cannot surf the city-wide labour market, so labour does not move with the ebb and flow of opportunity the way capital does in this unfair market.[17]

So it becomes apparent that, for a variety of reasons, including the lack of real 'free' mobility of capital and labour, market prices fail to provide a reliable measure of the real social value of goods or labour services. Such prices cannot, therefore, provide any sort of rational or just basis for determining wages paid to producers.

Individual contribution

There is a more fundamental problem for the neo-liberal idea of reward in proportion to the value of individual output. Even if we assume that market prices are accurate measures of the real social value of goods, we still have to be able to measure the precise contribution of particular individuals in generating such value: the component of the value of any particular product or service contributed by an individual. It is far from clear that this is possible.

No-one in a modern industrial economy produces anything on their own. All production is a complex collective effort involving the effective collaboration of dozens, hundreds and thousands of people, directly through the division of social labour within and between productive facilities and indirectly, through the use of knowledge, skills, tools and infrastructures provided by others in the past. And this is particularly the case within large corporate structures, which often integrate productive processes around the world.

Of course, special skills are still needed to allow individuals to participate effectively in achieving particular forms and levels of output. Yet, the production of such necessary skills is itself, typically, a similarly collective social effort. In this area also, monetary costs do not necessarily reflect real social costs of production. Individuals seldom bear the full monetary cost of the production of the skills they acquire: substantial public contributions are almost always involved.

A useful analogy here is that of constructing a machine mostly

made up of relatively easily built and cheap components. A couple of parts, necessary to complete the system, are more complex and expensive to produce. We need to have them in order to allow the machine to function. But this does not make the more costly parts more important or more productive than other parts in the overall operation of the system, nor mean that they contribute more to whatever task the machine is designed to perform.

It is true that if this machine is used to produce a particular output for sale, and we want to replace it when it wears out, then we must factor in the differing costs of the different parts into the total sale price of the total output of the machine, during its productive life. We can even attribute a proportion of the value of individual products to the replacement costs of the high-priced parts in question. But this does not mean the more costly parts are 'more productive' or that they actually, in any objective sense, contribute any such measurable component of value.

Suppose we add a new part to the machine and double the output of useful items from the same input of raw materials. This does not make the new part 'responsible' for half the new increased production. It is the addition of the new part to the other parts and their collective operation that are responsible for the total increased output. The new part does nothing on its own. It just so happens that it is needed alongside the old parts to create the new system. There is no reason why this new part need necessarily be complex or expensive.

Similarly, and despite claims to the contrary, it is not possible to trace particular components of value of the final output back to the productive contributions of particular individuals within social organisations of production. It might be claimed that we can see by how much the new employee – the new CEO, for example – has increased profits. At best, the increase is a result of new relationships between many individuals, collaborating and co-ordinating their particular activities. Perhaps the CEO 'brings out the best' in the work force, but so do the workers give their best in the changed circumstances – and give the CEO a chance to 'shine'. Perhaps the new boss terrorises the work force into increased intensity and duration of labour. Such increased profits could equally be related to new, cheaper or more productive technology, cheaper raw materials, or some other 'external' consideration, most obviously increased demand for the product.

This means that the distribution of such increased profits, and indeed of all revenues from the sale of products, whether they go to

workers, managers or shareholders, is not based upon any clear principle of distributive justice, of either reward for effort or reward for 'objective' contribution to the value of the product. Rather it is a matter of politics, of the differential power of the relevant parties, moderated by 'market' demands for the replacement of productive resources and competitive capital accumulation.

Working-class powers of collective bargaining through organised trade unions could be said to accurately reflect the intrinsically collective nature of productive activities. At the same time, such powers go some way toward the redressing the imbalance of power intrinsic in monopoly private ownership of productive resources and the capitalist wage contract. Yet high levels of unemployment since the 70s have assisted the employers to destroy trade union powers and institutions and practices of collective bargaining in favour of individual work contracts.

The neo-liberal ideology of reward in proportion to output of value thus stands revealed as little more than an ideological justification for radical income inequality. In particular, it serves to rationalise the exponential growth of executive salaries and the systematic attack by ruling groups around the world upon trade union rights and powers of collective bargaining.

Effort and sacrifice

Even if it were possible to precisely measure individual contribution to real social value of output, this could not serve as a fair or just basis for differential reward. If some can produce lots of good things with minimal time and effort, because of natural or social advantages (of special strength or special training), while others produce less with greater effort, it is far from fair to provide greater rewards for the more productive individuals.

They might have made efforts and sacrifices in the past in order to improve their productivity (though training and education need not necessarily involve much effort or sacrifice). Such effort and sacrifice may deserve appropriate reward, but any subsequent greater output as a result of such skills deserves no extra recompense.

In contrast to neo-liberal claims to the contrary, it seems to be generally true that individual effort is more readily assessable than individual contribution to the value of output. As Hahnel points out, individuals working together generally have quite a good idea about

who is working 'longer, harder, or at more dangerous, stressful or boring tasks', and who is slacking off, or doing easier or more enjoyable work.[18]

There are problems in devising workable systems of remuneration based upon such facts, problems of possible victimisation and 'tyranny of the majority'. And in line with the utilitarian arguments for equality of outcome considered earlier, clear limits must be placed on the nature and extent of reward for effort, so as to reduce inequalities of outcome.

It is arguably fair and just that all who can contribute should contribute equally, in terms of time and effort, according to their ability, to provide basic health, educational and social security benefits, along with basic subsistence, equally available to all, according to need. And benefits for extra work, over and above such basic contribution, cannot include the means to achieve any sort of privileged access to economic or political power.

In this context, a first approximation to an effort-based system would simply pay all a standard wage per hour worked. Such a system would produce smaller disparities of final outcome than one based upon the supposed value of individual output. Reward for effort, in combination with basic need satisfaction, provides a foundation for a fairer and more democratic system than the present one.

Shared responsibility and optimum working conditions

In the light of the above problems with the neo-liberal view, it could be argued that, allowing for provision of appropriate social welfare assistance for those unable to make valuable productive contributions, it is generally fair and just that individuals should be rewarded in proportion to their contributions of effort and sacrifice in the service of general social welfare. Such a system would be efficient and practical in encouraging optimum productive contributions from all concerned. Genuine effort and sacrifice in the acquisition of valuable skills would be rewarded, but such skilled labour would not thereafter receive any special reward.

However, this would still leave the distribution of social wealth, which is produced collectively, to be determined on an individualised basis. Disputes and conflicts over entitlements are likely to continue. The optimum outcome would be the creation of social structures and

relations that allow and encourage all to make the best possible contributions to the general social good. This would involve genuine sharing of difficult and demanding work, and genuine efforts to make all work as safe, intrinsically rewarding and empowering as possible. By all indications, this would require the abolition of capitalist market relations in favour of a socialist economy. Some of the issues about the possibility and feasibility of such a transformation were discussed in chapter 12 on Marxist legal theory.

Discussion topics

1. Does the market deliver fair wages? If not, why not?
2. Can wage inequality be justified in utilitarian terms?
3. What sort of legal interventions might improve the situation (that is, make it fairer)?

Additional resources

J. Wolf, *An Introduction to Political Philosophy*, Oxford University Press, Oxford, 1996.

Chapter 16

INHERITANCE AND RETURNS TO ASSETS

Inheritance

So far, we have considered only the neo-liberal defence of reward in proportion to individual labour contribution or output from labour as a basic principle of distributive justice. The neo-liberal defence of inheritance of private wealth, of monetary return to productive utilisation of assets and to share ownership, raises further difficult issues. Neo-liberals are staunch defenders of the right of property owners to freely hand on such property to others, including their descendants. But the institution of inheritance, in general, poses a problem for desert-based ideas of justice and fairness. Those who inherit wealth do not necessarily do anything to deserve it. Nor is there any reason to believe they are the 'right' people to make the 'best' use of it, in utilitarian terms.

In some cases of inheritance there will have been considerable productive effort involved. Person A might bequeath funds X to person B in recognition or appreciation of B's efforts as (unpaid) carer, looking after A or some other party. In this case, the inheritance could be identified as a reward for productive effort. Party B might have a legal claim on A's estate in these circumstances even if A has not actually willed them anything.

However, there is typically no legal necessity for any such productive effort or productive contribution as a condition of inheritance, no attempt to match a 'just' inheritance to the extent of such productive contribution. Some carers get far greater rewards, not on the basis of the quality or quantity of care, but rather the wealth of the cared-for

party. Most important for present purposes, there is no attempt by neo-liberals to demand any such requirement of service, or to defend the institution of inheritance in terms of such a requirement.

Rather, the neo-liberal defence centres upon ideas of private property ownership and the (fundamental) rights of the property owner. In particular, they argue that property ownership as reward must include the right of 'free' disposal of such property, including the right to 'gift' it to others, without outside interference. An individual has worked hard for many years to give their kids a reasonable start. What more admirable and unselfish reason could there be?

Utilitarian defence of inheritance

Here is what we might call the first utilitarian line of defence of inheritance. If the 'talented' people are not allowed to hand on their accumulated wealth to others then they will not do the best for society and we all lose out. So we should not deprive them of the opportunity to do so.

Once again, we are presupposing the unique efficiency of free markets, sustained by such people (neglecting monopolies, false needs, inequality and externalities). We are also assuming that the contributions of those concerned are necessary for the efficient running of the system, that they do their best only as long as inheritance is maintained.

But quite apart from such (highly problematic and unsupported) factual claims, it remains the case that inheritance undermines the fundamental principle of reward for effort and sacrifice. It also undermines any meaningful equality of opportunity. Those who hand monopoly control of scarce resources to their offspring – and the privileges and opportunities that go along with this – are handing poverty and powerlessness on to the rest of the population.

This brings us to what we might call the second utilitarian line of defence, explicitly spelt out by Milton and Rose Friedman, amongst others.

> The accumulation of physical capital – of factories, mines, office buildings, shopping centres, highways, railroads, airports, cars, trucks, planes, ships, dams, refineries, power plants, houses, refrigerators, washing machines and so on in endless variety – has played an essential role in economic growth. Without that accumulation the kind of economic growth that we have enjoyed could never have occurred. Without the maintenance of inherited capital the gains made by one generation would be dissipated by the next.[1]

The idea here is that the accumulation of capital – the expansion of the scale of productive organisation through reinvestment of profits – has allowed for the development and introduction of new technologies and achieved economies of scale that have increased productivity, created new products and reduced costs, to the benefit of all.

But, in the modern world, capital accumulation is not dependent upon continuity of private ownership. Public corporations can continue to organise such capital accumulation despite radical changes in share ownership. And privately owned corporations can easily be taken over and run by the state.

In some cases large organisations suffer dis-economies of scale, while modern technology allows for high efficiency in smaller-scale production. Large organisations can also be associated with substantial political power and monopoly pricing. Breaking down concentrated private capital accumulations – with the death of the owners – could be highly desirable in order to increase efficiency and democracy, and reduce prices to consumers.

This leads to the final utilitarian defence of inheritance, to the effect that the children of the rich are particularly well suited to make the best use of their parents' accumulated wealth, to the general good of all.

This implies that such offspring have some special skills, not available to others in the population. But, in the first instance, it is far from certain that any such skills actually exist at all. Arguably, running a big corporation requires no such special skills, especially given the resources of specialised assistance available to owners and controllers of larger operations. In such a context, relevant knowledge and skills can be supplied by others.

Even if such skills did exist, there is no evidence of their 'genetic transmission' within the ruling elite. They would not be the kind of thing directly encoded in the genes, but would inevitably require substantial social development. And even if there were such genetic transmission, (a) there would still be an issue of the distribution of such skills outside the ranks of the ruling elite; others could still have inherited more of them, but be prevented from making good use of them by inheritance of productive resources within the ruling class; and (b) there is no reason why such genetic privilege should transmute into the social privilege of private wealth.

If such skills do exist, they will themselves be products of social privilege, of access to special educational and other advantages. The

children of the wealthy elite might indeed have such special access. But this in no way justifies the further privilege of inheritance of vast private wealth. Two wrongs do not make a right.

This leaves only the idea that, given they have been allowed to develop such skills, it would be inefficient not to make the best use of such skills. But (a) it is far from clear that the offspring of the wealthy have to inherit such private wealth in order to make proper use of such talents; (b) if they do demand such vast rewards, some such wealth could easily be redirected to give such skills to other less greedy people; (c) it is far from clear that such talents, and the kind of corporate capitalism they sustain, are really the kinds of things we want to maintain in order to maximise social welfare. Arguably, a planned socialist economy is much more defensible as a means to achieving such welfare. The utilitarian defence of inheritance is thus seen to beg a great many questions.

It would be easy, in theory, to abolish all such inheritance, with accumulated business assets falling into public ownership upon the death of the owner. Such assets could then be utilised on the basis of democratic decisions concerning genuine social welfare priorities. They could contribute to increasing the chances of genuine equality of opportunity by providing increased educational and medical resources for those with special needs.[2]

Productive assets

Turning to the issue of rewarding the use of productive assets, it is first important to distinguish different kinds of productive resources. The strength, skill, knowledge and intelligence of human beings can function as productive resources (what Marx called 'labour power') alongside 'external' productive forces of land, raw materials, tools, machines, infrastructures and knowledge stored in books and databases.

Both sorts of assets could be acquired through inheritance: some, at least of the former, through genetic inheritance of parental characteristics, the latter through inheritance of property. In the latter case, the same issues apply as with (differential) inheritance of any goods; because there is no necessary effort or productive contribution involved in acquiring such goods, a theory of reward for effort and sacrifice – or productive contribution – suggests that individuals should not derive any reward from them.

In the former case, there is no question of depriving individuals of such bodily and mental assets. But again, it is clearly unfair that those with particular genetic advantages should thereby gain further economic advantage. They should still be rewarded in proportion to the effort and sacrifice they put into useful productive labour, compared to others.

Effort and sacrifice are generally needed to further develop and apply socially useful bodily or mental talents. Arguably such effort should itself be appropriately rewarded. But once acquired, by whatever means, such assets should not command any extra reward, merely the standard reward for effort and sacrifice in their further use.

Sacrifice

Neo-liberal theory demands that individuals be able to acquire external productive assets as rewards for productive effort (or through inheritance), and then benefit from possession of such assets, without further effort, through renting them out to others or hiring others (for a wage, as both managers and workers) to use such assets to produce a surplus. In this case, huge rewards can be achieved with no effort whatsoever, with others paid from returns to the assets in question to supervise all such processes, of hiring, firing, supervising, investing, etc.

As the idle rich become richer they simply hire the best 'help' in managing, and further increasing, their assets. If they so choose, they can sit back and do absolutely nothing as the wealth flows in – merely consuming expensive luxuries.

The neo-liberal 'ethical' claim is that there is still 'sacrifice' involved, and therefore reward deserved. For in choosing to make such assets available for productive utilisation (including money stored in interest-bearing accounts in banks), creating jobs and useful products (providing rental accommodation, etc) the owners have 'sacrificed' the consumer goods they could have acquired by selling the assets in question (or spending the money) outright.

This argument can be seen to run parallel to the neoclassical economic idea of wages as reward for leisure time foregone. The worker freely chooses to trade off the pleasures of leisure time for the pleasures of money to spend (acquired through intrinsically unpleasant labour). That particular idea has little factual basis, given that most workers have to work to live, and not all work is, or need be, intrinsically unpleasant.

The factual – and moral – foundations of the argument about foregone consumption by the wealthy are equally shaky. It is difficult to see how there could be 'sacrifice' of consumption if such use of assets actually generates substantial increased unearned income for future consumption. This is the opposite of sacrifice.

Neo-liberals attempt to strongly differentiate between sacrifice of consumption today – contributing to socially useful productive output – and consumption tomorrow as legitimate reward for deferred gratification. But it is far from clear that the majority of those members of contemporary capitalist society who do own productive assets (or wealth sufficient to sustain significant productive investment) do, through productive use of such assets (or investment), actually suffer any significant material deprivations. On the contrary, their living standards typically remain far higher than those of the property-less majority of the population.

In reality, of course, it is seldom a question of 'today' or 'tomorrow' but really just an ongoing feedback of profit generation, reinvestment and luxury living, from before the birth of any particular member of the privileged elite until after their death. The sacrifice argument turns out to be a 'red herring' fallacy – of the kind considered in chapter 3.

Risk

Here, neo-liberals fall back upon the idea of 'risk'. There is always some risk – in the sense of possibility of loss – involved in productive use of assets, as opposed to sale and direct consumption of the proceeds. Future benefits are much less certain than current ones.

The degree of risk is in practice dependent on the scale of assets involved. Small businesses are, indeed, at constant risk of going under, because of the economies of scale and monopoly power of bigger businesses with whom they compete and from whom they purchase necessary inputs, because of changing public tastes and the ups and down of the business cycle. Success in this world is, indeed, largely a matter of chance, with few succeeding while many go under. Most important, it is by no means necessarily hard work and virtue that are rewarded by the vagaries of competitive market forces.

For bigger businesses, by contrast, there is less risk involved. Wealthy owners can avoid risk through employing clever lawyers, accountants, marketers and others to help cut costs and taxes and

boost profits. They can insure against problems and take advantage of government assistance, cheap labour overseas and monopoly power in ways not available to smaller operators. Only major world recessions challenge large operations. Owners of such operations will typically have plenty of resources 'salted away' for such eventualities. Even if they lose all productive assets they are still far ahead of the majority of the population.

Neo-liberals see the choice not to sell productive assets – and consume the proceeds – as justifying the unearned income from such assets. But, often it may be more profitable to postpone sale of assets until market conditions improve. Here it is immediate sale, not investment or use that incurs the loss. And sale of substantial (including inherited) assets can generate wealth far beyond anyone's practical abilities of immediate consumption.

Moreover, just because assets are sold, this does not mean they necessarily cease to be used productively. Typically others will buy them in order to continue to 'operate' them to make a profit. Just because they are not sold does not mean they are put to their best use, or, indeed, that they are contributing anything to the general social good. They could be producing and circulating weapons, asbestos, chlorine, cigarettes or unnecessary and addictive drugs. So utilitarian arguments for social benefits of investment or 'use' rather than sale are not at all strong.

Stockholders' risks

With shareholding also, the right-liberal argument is that those who purchase and hold onto stock rather than purchase items for immediate consumption are thereby suffering deprivation and taking a risk – of total loss – which justifies the unearned income from such shares. Again, this neglects a similar 'two worlds' situation of large and small investors. The 'mum and dad' investors are, indeed, exposed to great risk in what is, in effect, a lottery. They are gamblers – albeit unwillingly so in respect of participation in compulsory pension funds – where others gamble with their future subsistence needs. And gambling is not generally recognised as a very ethically elevated pursuit. Many lose their (sometimes hard-earned) resources and turn to others for support.

On the other hand, those with significant sums to invest do not expose themselves to significant deprivation even if particular invest-

ments happen to founder. This is particularly so, given the limited liability of joint stock companies in the modern world, and the diversified portfolios and 'inside knowledge' of the wealthy. If investors know what they're doing (including having access to 'insider' knowledge of impending developments), or, even more so, if they hold shares in the operations of big corporations with substantial power to control prices, markets and governments, there is very little 'risk' in the sense of possibility of significant loss. In this sort of case they would seem to be rewarded for not actually taking risks (and for contributing to price fixing, anti-competitive behaviour and insider trading). The monopoly profits of the big corporations in effect transfer the new wealth created in the competitive sector (of smaller businesses), along with the surplus value created by their own workers) to the pockets of the idle rich.

If investors do not know what they are doing, or do indeed, choose to gamble on genuinely high-risk investments, then they would seem to be rewarded (when they happen to 'win') for irresponsibility. Perhaps, in some few cases, the high-risk ventures do have significant potential for social good. If so, the fact that they are 'high-risk', that they stand little chance of survival, is a serious indictment of the current market system.

Current legal arrangements treat stockholders as 'owners' of public corporations. Only a handful of families – less than 10 per cent – own about half of all stocks in public companies worldwide. It is often said that such stockholders provide a vital public service in funding such corporations' productive activities through their stock purchases. This is the core utilitarian defence of the system that rewards such stockholders with substantial yearly dividends. But, as Marjorie Kelly points out, new sales of common stock are rare events:

> Among the Dow Jones industrials, only a handful have sold any new common stock in 30 years … in recent years about 1 in 100 dollars trading in public markets has been reaching corporations.[3]

The rest is going to those who have bought shares simply to garner ongoing unearned income, and to speculators: gamblers aiming to realise their capital gains or minimise their losses.

In fact, as Kelly points out, the situation is much worse than this. If we consider the increase in value of exchange-listed stocks in the United States from 1998 to 1999 the proportion attributable to sales of new common stocks was around 7%. But this was offset by stock

buy-backs far greater than new equity issues, to the tune of an extra $US 267 billion. As Kelly concludes:

> The stock market, in reality, is not 1% or 7% productive, it is less than 0% productive. And that's not even including dividends, which in 1998 extracted an additional $238 billion from corporations … when you look back over two decades – 1980–2000 – you can't find any net stockholder money going in – it's all going out … Rather than capitalising companies, the stockmarket has been de-capitalising them.[4]

In theory, stockholders are entitled to all company profits – or revenue above costs – even though they have done nothing to produce this surplus. In practice, they receive about one-third of it as dividends, with the other two-thirds used by the corporation, but still 'booked in the balance sheet as stockholders' equity'.[5] With increased profits, both dividends and share values increase. And, in general, 'equity grows year after year, while stockholders never contribute another cent'.[6]

So if someone actually did earn the money to purchase common stock (through some genuine effort and sacrifice) and did purchase it from the company at the time of issue, and did hold onto it thereafter, rather than selling it for capital gain, they are still receiving an 'infinite payback from a onetime hit of money'.[7] The workers who create this new wealth on an ongoing basis through their effort and sacrifice see none of it and have no power of corporate governance, with boards of directors elected purely by shareholders and legally required to govern purely in shareholders' interests, which means profit maximisation above all else. As Kelly says, one relatively 'tiny contribution' allows stockholders to 'install a pipeline and dictate that the corporation's sole purpose is to funnel wealth into it'.[8]

As we have seen, the great majority of those shareholders exercising such powers and reaping such benefits from this endless and generally expanding flow of unearned income have not actually contributed anything to the corporation in question. Their 'investment' dollars have actually gone to other speculators (or rentiers) from whom they have purchased the shares.

> While we still call stockholders the owners of major public firms, they do not, for the most part, manage, fund or accept liability for 'their' companies. Ownership functions have shrunk to virtually one dimension – extracting wealth.[9]

It would be easy, in theory, to change all of this, with democratic

representation of employees, clients and communities in corporate governance. Profits could be shared between those contributing to their production or those able to derive most real benefit. And the law could require that genuinely fair wages and the health and safety of workers, communities and the natural environment come before profit maximisation.

But the 10 per cent of families around the world who own half of all stocks, including big chunks of pension accounts, are in a strong position to prevent this from happening. Arguably, their control of mass communications media is crucial in fostering continued acceptance of this ridiculous and immoral situation. Kelly comments that in medieval times the mass of people acquiesced in an established order that accorded half of society's income above subsistence to 1–2 per cent of the population 'out of religious awe'. More specifically, the Church sanctioned the claim that the feudal nobility and the upper Church hierarchy 'deserved' and 'required' such surplus on account of their active role in physical and spiritual 'protection' of the general population.

'Today we acquiesce out of financial awe, believing the wealth of the few is a natural consequence of economic forces too complex for ordinary mortals to comprehend.'[10] We assume that these people are somehow contributing to the 'general good' even if we cannot see how.

Corporate governance

The principle limits to shareholder power come not from government or from any kind of democratic community decision-making, but rather from executive authority within the corporation itself. Supposedly, directors and CEOs are employees of the shareholders, with boards of directors elected by shareholders. To the extent that CEOs fail in their legal ('fiduciary') responsibility of maximisation of profit and return to shareholders, wealthy shareholders can take legal action against them or pressure boards to sack them.

In reality, boards of directors are generally 'hand picked by the CEO and the previous board, and rubber-stamped by shareholders'.[11] And corporate structures are bureaucratic hierarchies, with information and decision-making power concentrated in the hands of such directors and CEOs, and other employees caught up in 'vertical relations of domination and subordination'.

As long as those at the top have appeared to be delivering the desired result, as long as profits and dividends have continued to increase (recently achieved through sackings, overseas sweatshops, corporate welfare, tax avoidance and 'accounting sleight of hand'), shareholders have been prepared to reward them with exponentially increasing salaries, bonuses and stock options increasing in value. 'In 1991 the standard big company CEO in the United States earned 140 times the pay of the average worker; the multiple is now nearer 500 times.'[12]

At the same time, CEOs positions of power have allowed them to further enrich themselves even in face of losses, crime and incompetence, at least in the short to medium term. CEOs and other executives are always the first to know when the game is up, unload their blocks of shares and salt away the proceeds safe from confiscation before share values plummet. As Beder notes: 'Altogether 29 [Enron] executives cashed in $1.1 billion of stock from 1999 to mid-2001', prior to the share-price collapse. The company founder, Kenneth Lay, was selling Enron shares worth $100 million in 2001 while assuring employees and other shareholders that Enron 'was continuing to have strong growth'. And senior executives paid themselves $55 million in bonuses days before the company filed for bankruptcy – and stock prices plunged to less than a dollar each.[13]

The principal victims of such losses are increasingly ordinary workers, reliant upon stock market super-funds for retirement income, while the diversified portfolios and sheer bulk of the assets of the super-wealthy offer substantial protection. In the United States in 2002

> the average CEO compensation package equalled $10.83 million according to the New York Times … While stockholders – including workers who depend on the stock market for their retirement savings and pensions have lost $7 trillion since the stock market peak … median CEO pay increased by 6 % in 2002.[14]

Similarly, in Australia, while the profits of the top 500-listed companies were down by an average of 32% in the six months to 31 December 2001, executive salaries rose around 7% across the board. Chief executives now earn an average of $1.4 million, compared to an average wage of around $40 000 and an aged pension of about $200 per week.

> In 1976 CEOs earned three times the average wage. Today, the average CEO, taking executive share options and bonuses into account, is paid 30 times the average wage.[15]

Nor do these people apparently have any qualms about the situation. A poll of *Fortune 1000* CEOs in 2002 'revealed that 87% thought their pay was just what they deserved, 11% thought they were underpaid, and only 2% confessed to feeling over-rewarded'.[16]

In the longer term, wealthy stockholders can, on occasion, exert power through hostile corporate takeovers and firing of CEOs where profits are deemed inadequate. 'Activist' boards sacked the CEOs of 24 huge US corporations, including General Motors, IBM, American Express, Kodak and Westinghouse between 1991 and 1993. So can such shareholders pressure government to increase the legal protection of their wealth, shielding them from the fraudulent activities of corporate executives.

The situation became particularly clear following the scandals of Enron, Arthur Anderson, WorldCom and others in the United States. Things were fine so long as the fraudulent actions of CEOs and their accomplices continued to boost return to shareholders, no matter that Enron 'bought out government, dismantled regulatory infrastructures, destroyed overseas villages to build power plants, evaded taxes and manipulated energy prices'.[17] It was fine for WorldCom CFO Scott Sullivan to 'devise clever ways to remove debt from balance sheets'[18] so long as profits continued to rise. But once shareholders' returns were seriously threatened through such fraudulent accounting and insider enrichment, powerful political and legal forces were mobilised to ensure shareholder protection, with the errant CEOs brought to book.

It is far from clear that these measures will be effective in reining in CEOs' greed. It seems that corporate executives, with the help of clever accountants and lawyers, will find new ways of perpetrating fraud on a similar scale in the future, bypassing half-hearted law reforms. For present purposes the important point is not so much the effectiveness of such measures in achieving their stated aims. Rather, it is the fact that political authorities aim to achieve only shareholder protection, with no consideration of proper protection of the wider community, of the environment and of democratic rights from accelerated corporate devastation.

Power

The greatest problem with private ownership or effective control of productive assets is not that of the financial benefits that flow to such

owners and controllers per se, but rather the benefits of effective social power that accompany such material rewards.

In modern society, returns to assets – and senior managerial salaries in big corporations – represent a qualitatively different order of income from that available to the great majority of wage and salary earners. Whereas the latter allow for little beyond the most basic subsistence, the former provide a huge pool of surplus wealth that can be used to create new jobs or destroy jobs, transform the physical world, take over and run major means of mass communications and thereby profoundly influence public ideas and perceptions, finance the campaigns of political parties and shape their policies when they are in power.

Even if recipients of this surplus wealth confined their expenditure to the enjoyment of the benefits of consumption, without directly seeking any such social power and influence, such wealth allows a scale of consumption that profoundly influences the social division of labour and use of productive resources. In particular, it means a substantial concentration of productive effort and resources in the provision of expensive luxuries, at the expense of provision of necessities for the majority of the population, including high-quality public health, education and welfare benefits. Albert notes that

> Having a few members of society own [or control productive assets], decide their use and dispose of the output and revenue they generate has meant that this privileged group has always had more wealth, more income and more economic power than others in society.[19]

And economic power has typically been the foundation for political, legal and ideological power – power to shape the life and health and ideas of the whole population.

As Albert says, our most basic ethical intuitions tell us that – if possible – no-one should have any disproportionately greater power or income by virtue of owning or controlling means of production. And the obvious 'logical step' we can take to accomplish this is 'to simply remove ownership [and individual control] of the means of production from the economic picture'.[20]

Justice and fairness demand, as Albert says, that

> no-one has any ownership [or control] of means of production that accrues to him or her any rights, any responsibilities, any wealth, or any income different from what the rest of the economy warrants for him or her [through his or her actual productive contribution or 'fair'

share of welfare support]. No-one has wealth, income or economic influence different than what anyone else has due to having different ownership [or control] of means of production.[21]

Discussion topics

1. Is it fair that some should inherit great wealth while others inherit nothing?
2. Can inheritance be justified in utilitarian terms?
3. Is it fair that individuals should derive income by virtue of ownership of productive property? If not, why not?
4. Can private ownership of productive assets be justified in utilitarian terms?
5. What sorts of legal interventions might improve the current situation in terms of justice and fairness?
6. Briefly outline the role of the current tax system in promoting inequality. How should the system be reformed?

Additional resources

E. Sangkuhl, 'An Inequitable, Complex and Inefficient Taxation System' (2003) 28 *Alternative Law Journal* 225–29.

SECTION THREE

Law and contemporary social problems

Chapter 17

TORT LAW REFORM

Torts

Classical liberal theory allows total 'freedom' of individual use of and benefit from private property only up to the point at which such freedom impinges upon the corresponding freedom of other private property holders. At this point, those others have a right to take legal action to protect their property and themselves.

The civil law, as government-assisted self-help, makes sense from a liberal perspective. In particular, tort law provides the means for individual victims to help themselves in seeking monetary compensation for encroachment, pollution, injury or property damage. It has been argued that such tort law 'rebalances' a system undermined by externalities, bringing economic costs and benefits back into line with social costs and benefits.

It is certainly true that tort law can compensate victims for their suffering as a result of cost externalisation by others. It can offer the victims of pollution, work injury and sale of toxic products by business a means to gain some redress. But the tort system is incapable of fully compensating for the externalities of capitalist market relations. The problems of tort law are the very same, central problems of the Coase theorem, considered earlier: imperfect information, unequal power and burgeoning transaction costs.

Tort litigation is extremely expensive and, therefore, without comprehensive legal aid, beyond the reach of ordinary people even in the industrialised first world. Libertarians in government have themselves led the way in slashing legal aid funds, thereby leaving only the

vagaries of limited legal charity or a guaranteed big cut of easily obtained and substantial payouts for lawyers as means of access to tort law for the majority of the population.

Many, typically poorer, people will not even know that they are victims of others' negligence, because they are poorly informed about risks and dangers. They will perhaps know that they are suffering, but not why they are suffering – something that can itself require access to substantial resources to discover, for example, access to pollution monitoring equipment, complex medical tests, etc.

Very limited resources are available to the majority of victims of the sale of dangerous products and pollution by big corporations to mount tort actions. The resources available to bigger corporate perpetrators to defend, and discourage, tort actions – or string out proceedings to bankrupt opponents – are virtually unlimited, and can be written off to tax or passed on to consumers or shareholders. This is particularly relevant in the light of the scientific complications involved in establishing the facts in such cases.

All the evidence suggests that the great majority of victims of (serious) medical negligence take no action. Dewees, Duff and Trebilcock estimate that currently, in the United States, 'only about one in eight victims [of medical negligence] initiates a claim and only one in sixteen … receives any payment'.[1] Similarly in Australia, 'only a small number of the original pool of potential claims make it to court'.[2] And, in part due to massive difficulties in proving fault and causation, 'patients who [do] litigate generally do not win if the claim goes to hearing'.[3]

The evidence similarly suggests that only a tiny percentage of victims of product-related injuries take legal action. Data from US asbestos claims shows that 'the average total compensation per closed claim was $60 000 and after deducting plaintiff's legal fees and other expenses net compensation received by plaintiffs averaged $35 000 or 37 per cent of total expenditures by defendants and their insurers'. As Dewees et al argue, 'this amount … reflects a tendency to under-compensate fatal injuries and under-compensate for economic losses in the case of serious injuries'.[4]

Air and water pollution causing increasing rates of particular diseases seldom give rise to tort litigation. Because of difficulties in establishing a direct line of causation from specific discharges to specific injuries, polluters escape liability, despite causing harm on a large scale.

Tort action is futile if the defendant lacks resources – or insurance – allowing them to pay compensation. Small businesses might be bankrupted by big claims, while companies – large and small – can choose 'strategic' bankruptcy and transfer of resources out of the organisation, to avoid payment. As Abel notes, the poor, as more likely victims of crime, are also likely to be victimised by tortfeasors who cannot or will not pay compensation.[5] There is no guarantee that plaintiffs' interests are well served in out of court settlements. Such settlements deprive society of the public disclosure of misdeeds.

The pattern of third party liability insurance delimits the scope of tort claims. The system comes under strain when crime, greed and incompetence in the insurance industry lead to failure of coverage and increasing premiums. And wealthy insurance companies employ big city law firms, adept at presenting poorer claimants as liars and malingerers.

Tort damages themselves perpetuate inequality insofar as those with greater wealth, higher wages or earning potential, greater 'reputation', better health and longer lifespan (by virtue of social privilege) receive correspondingly higher payments. Continuing high levels of accidents and injuries, with 400 000 injuries needing hospitalisation and 8000 'accidental' deaths in Australia each year,[6] testify to the failure of the system as a deterrent mechanism. There are major deterrence problems even with 'successful' litigation. As Abel notes, tort law

> consistently violates the basic principle of proportionality between the wrongfulness of defendants' conduct and the magnitude of the penalty imposed. Because punishment is a function of the harm caused, it is either too severe or too lenient. It is too severe where momentary inadvertence results in a catastrophic injury … too lenient where egregiously unsafe conduct happens to cause little or no injury – by chance or through the intervention of others, eg, a negligently constructed and maintained office building consumed by fire in the middle of the night when it is empty.[7]

As Coasian ideas indicate (see chapter 13), the wealthy and the powerful can find it cheaper to pay the tort costs of harming other people, rather than taking action to prevent such harm in the first place. Wealthy perpetrators can pay to avoid court action (through intimidation and out of court settlements), avoid public airing of the issues (creating pressures for possible regulatory reform), reduce costs and maintain their 'good name'. And they can afford comprehensive insurance to protect themselves from the consequences of their actions.

Issues of radical resource imbalance between different class groups and the problems of establishing causation in relation to pollution and toxic product litigation undermine the deterrent effects of tort law in relation to much corporate wrongdoing. In addition, a major motivation for relocation of manufacturing operations to poor countries has been to escape any sort of effective legal regulation or liability for such cost externalisation. And this, in turn, puts further pressure on such liability in first world countries, struggling to hold on to or attract mobile capital.

Nineteenth-century common lawyers, strongly influenced by libertarian ideas, developed principles of contributory negligence (of victims' failure to meet a required standard of care in their own protection) and voluntary assumption of risk – including tacit acceptance of known dangers – to offset the legal responsibilities of those creating, overseeing or inflicting dangerous situations. As Fleming says, this latter approach

> was most ruthlessly invoked in employment cases so as to debar injured workers on the barest finding they continued in their job after becoming cognisant that the working conditions were hazardous … with the growing strength of industry and changing social ideas, this draconian doctrine began to yield … culminating in the drastic reformulation of the defence in the great case of *Smith v Baker*.[8]

Libertarians or neo-liberals in power have not necessarily defended the tort system. The more such a system has been used against big business and powerful professional groups, particularly doctors, the more business and professional groups have lobbied for its restriction in favour of 'alternative dispute resolution' mechanisms where the dice may be even further weighted on the side of wealth and power (with no public airing of issues to serve as a basis for regulatory legislation, no system of precedent, no appeal, etc). Governments have responded by winding back the rights of ordinary people in this area, bringing back ideas of voluntary assumption of risk and capping compensation payments.

Reforms

In Australia, developments in the medical area, in particular, have driven major legislative changes to tort law. Supposedly, there was a crisis in medical indemnity, associated with an explosion in the number and cost of claims against medical practitioners. This was par-

alleled by sharp rises in public liability insurance premiums, threatening the viability of small businesses and community organisations. Doctors pointed to the Calandre Simpson case, involving an award of $14 million to a woman who suffered cerebral palsy as a result of a negligent forceps delivery. The case supposedly led to upwards adjustments of claims across the board, and to lawyers' 'no win, no fee' advertising encouraging increasing litigation. They also blame the rejection of the *Bolam* test by the High Court in *Rogers v Whittaker* (1992) 175 CLR 479. *Bolam* (1957) 1 WLR 582; (1957) All ER 118 (QBD) had provided a defence for doctors in relation to provision of information to patients by preventing any action supported by a 'responsible body of opinion within the medical profession' from being found to be negligent. In *Rogers v Whittaker*, the High Court decided that it for the courts to decide what constitutes 'reasonable care' in such cases.[9]

As Skene points out, 'her wealthy family background' and the seriousness of her injuries played a significant role in making Calandre Simpson's payout so substantial, with the great majority of claims being much smaller. Skene also casts doubt upon suggestions that 'no win, no fee' advertising or the rejection of *Bolam* contributed significantly to an indemnity crisis. Few applications for 'no win, no fee' are accepted by law firms (presumably only 'sure things'), and patients still have to pay for experts' reports, court costs and defendants' legal costs if they lose. Similarly, 'it is conceivable that a court might find a doctor had been negligent even if he or she acted in accordance with accepted medical practice, but such a finding would be most unusual'.[10]

Furthermore, it is probably the case that the number of claims

> has not increased more than would be expected in view of the increase in the number of medical services provided under Medicare [up 66% in 15 years] and in hospitals [up 76% in 15 years] and the increased range of services. Doctors can now test pre-natally for congenital and genetic abnormalities and detect cancers at an early stage when effective intervention is possible. If these tests fail, people want to sue.[11]

Similarly, improved medical techniques now allow individuals to live for many years where previously they would have died. And the costs of their care can be very substantial.

At the same time, as Skene points out, after 1999, there was a 'remarkable turnaround' in favour of defendants in High Court appeals involving personal injury claims, from 80% pro-plaintiff

between 1987 and 1999, to 22% in 2000. She suggests that such decisions reflect a new concern that 'people should take greater responsibility for their own action'.[12] In other words, the courts were moving to restrict liability prior to legislative intervention.

Nonetheless, the concerns of doctors, small businesses and community organisations facing rising insurance premiums, and local councils, hit with some substantial payouts to victims claiming a failure to warn of geographical hazards (to swimming and diving in particular), seem to have significantly influenced legislators. As Skene points out, the Ipp Committee, a federal government panel of 'eminent persons' set up to review the law of negligence,

> recommended that a modified form of the *Bolam* principle should be adopted and this change has already been enacted in New South Wales, Queensland, and Tasmania and is included in a Bill before the Victorian Parliament. [Now] in New South Wales, a 'person practising a profession' is not negligent 'if it is established that [he or she] acted in a manner that [at the time the service as provided] was widely accepted in Australia by peer professional opinion as competent professional practice'.[13]

As Booth and Varghese point out:

> The [Ipp] panel's Terms of Reference asserted that the 'award of damages for personal injury has become unaffordable and unsustainable as the principal source of compensation for those injured through the fault of another'. Not surprisingly, the eventual recommendations of the Ipp Report were designed specifically to reduce the number of claims being made.[14]

Before the Ipp Committee was convened, in New South Wales, the *Health Care Liability Act 2001* had already limited the amount of compensation that could be awarded for personal injury claims. Subsequently, following Ipp's recommendations, legislation in a number of states has reduced the limitation period for adults commencing personal injury claims from six to three years, capped damages for non-economic loss at between $250 000 and $371 000, and banned exemplary, punitive or aggravated damages, as well as limited legal costs recoverable by plaintiffs from defendants and reduced damages for loss of earnings. The reforms introduced thresholds of injury necessary before a general damages claim can be made. In New South Wales, 'damages will not be awarded for non-economic loss unless the injuries are assessed at more than 15% of the most extreme case'.[15]

These thresholds, along with measures of the extent of injury to serve as a basis for compensation, are calculated on the basis of American Medical Association guides. As Booth and Varghese point out, they are arbitrary and unclear in many areas. There are problems also in their use by panels of doctors to assess the extent of injury involved.

> Plaintiff lawyers already have a difficult time finding doctors prepared to criticise the performance of their peers. Under the new process, assessments will ultimately be performed by medical panels, with no rights of appeal. Given the pervasive anti-litigation attitude of the medical profession, bias and the perception of bias is inevitable when medical panels assess injuries in medical negligence claims.[16]

In the medical area, mediation has become mandatory in New South Wales and Victoria. Such mediation can produce acceptable results only where the parties are equally powerful and well informed, not in situations of radical inequality.

Health complaints bodies offer cheap and accessible means for some victims to gain compensation and disciplinary action against negligent practitioners, including the use of a confidential report by an independent medical witness that might offer some hope of objective assessment without further legal implications. And a shift to 'structured settlements', whereby 'payments are made over the whole of the plaintiff's life, rather than in a lump sum',[17] which can be lost through fraud or unwise investment, does have potential advantages.

But overall, there seems to little on the positive side of the reforms, and much that is negative. As Booth and Varghese conclude, the changes

> further reduce the discretion of courts to determine cases on their merits … they erect more and higher hurdles for plaintiffs to jump before they can recover compensation for personal injuries to which they should be entitled … Despite the alleged and oft-quoted 'insurance crisis', insurers have returned to massive profits recently. But the cost … of medical negligence … will continue to be borne by patients for many years to come.[18]

An improved system

It would be possible to create a more egalitarian and effective system of compensation and deterrence. This would include comprehensive public legal education and public legal services with equal access by

all to high-quality legal advice and representation, at no direct cost or minimum cost. The former would help overcome problems of victims who fail to take action through ignorance. As Abel says, 'lawyers, the legal system, cultural norms and support groups' should provide strong encouragement for victims to claim redress. The latter would address problems of victims who fail to take action through lack of resources and fear of debt and destitution through legal costs.

The imbalance of power and resources between major corporations and government departments on the one side and individual workers on the other could be addressed through denying the former the opportunity for any legal spending beyond that provided to the latter by the public system, or by ensuring that the former provide the latter with funding to match their own extra spending.

Similarly, compensation should be the same for all – aiming to provide decent living conditions – with no extra payments for private wealth lost or higher income forgone. Or those on lower incomes could be compensated by higher payments and those with private wealth left to look after themselves. Abel argues that 'those who enjoy the privileges of wealth and income should bear the burdens of loss and loss insurance'.[19]

There would have to be broad community input into assessment of proper standards of care for particular professions. A responsible judiciary should take an active role in ongoing determination of the rationality, morality and competence of particular professional practices, with a strong emphasis upon rights of properly informed choice in all areas.

Causation problems could be addressed, in part, by extending something like the market share principle to the likelihood of a defendant's culpability in particular cases of injury (as considered in chapter 5). Thus, it should not be necessary to prove a direct causal link between, for example, this discharge and this victim's illness. Rather, the victim should be compensated in proportion to the increased risk created for the affected population by the discharge in question.[20]

All those who benefit from activities potentially exposing others to risk of harm must maintain insurance cover sufficient for effective compensation of victims. Payment by those found liable should be proportional both to the seriousness of the offence and the perpetrators' ability to pay (as measured by asset values), with extra payments, beyond victims' compensation, going to fund the legal system and public welfare provisions.

The criminal law should take a more active role in prosecuting risky and dangerous behaviours likely to produce accidents and injuries, particularly in areas of work safety, dangerous products and pollution. Criminal and civil proceedings could be combined in cases of actual injury, with civil standards of proof applied in cases of serious corporate crime, to facilitate prosecution and deterrence.

Strong measures would be instituted to prevent corporate perpetrators avoiding payouts, not covered by insurance, through strategic bankruptcies or drawing out the legal process to demoralise or bankrupt plaintiffs. From the start of legal action, assets of companies, subsidiaries, directors and their relatives must be effectively frozen, to the value of likely damages.

Secret out of court settlements should be banned, in order to prevent powerful offenders from manipulating and pressuring victims and avoiding public scrutiny. Broad and detailed public disclosure should become a central feature of tort actions. As Abel argues, tortfeasors must be made to fully acknowledge the extent and nature of their wrongdoing and apologise to their victims.[21]

No-fault schemes

A modified system would still absorb substantial resources in preparing and running cases, which could be better used compensating victims and reducing the likelihood of injuries in the future. It would deal only with injuries that can be proved to be caused by others' negligent action, thereby neglecting the substantial number that cannot.

Comprehensive state-funded no-fault compensation schemes for accident and injury could address these points. Dewes et al argue that such schemes in New Zealand and Sweden 'provide compensation [for road and medical injuries] faster, more effectively and more widely' than the tort system, 'with cost savings of 10% to 24%'.[22]

Such systems (including Australian state and territory compensation schemes for workplace injuries) have been rightly criticised on grounds of inadequate payments, failure of deterrence (with the public, rather than perpetrators, paying for compensation), and failure to address disability and loss of income not attributable to accident and injury.

Other public welfare systems exist in some jurisdictions to address some of these other causes of income loss. But in light of the inade-

quacy of such systems, Sugarman and others call for broad compensation schemes focused upon all who lack income through unemployment, regardless of causation.

Sugarman proposes 'mandatory short-term sick leave … up to six months, payable regardless of whether the disability is work-related' and substantial expansion of the social security system to provide ongoing payments of two-thirds of pre-tax income up to twice the national average wage for the wholly or partially disabled and the long-term unemployed, financed through a payroll tax. 'Employers would be required to provide enriched health care programs, including rehabilitation costs to employees and their families, while other individuals would receive similar benefits financed through general tax revenues.'[23]

Such proposals for general expansion of social welfare benefits have been criticised on grounds of cost and 'moral hazard', as well as failure of deterrence. It is suggested that workers would abuse a system with high levels of income protection for all those unemployed. Sugarman responds by setting benefits at a lower level for the healthy unemployed than the disabled. But it should be possible to provide a genuine living wage for all unemployed people, healthy and injured, with higher payments for all workers.

Major corporations have the greatest potential for inflicting large-scale harm through inadequate health and safety in the workplace, mass marketing and sale of dangerous products and discharge of toxins. The monopoly power of these corporations puts increased pressure upon smaller businesses to cut corners in these areas also.

It is the directors and major shareholders of the big oil and car companies who have done most to develop and maintain an inherently dangerous car-based transport system, through pressure upon consumers and governments to prevent the development of safer, more efficient public transport alternatives. It is these people – and the politicians they fund – who continue to benefit from such a system, with working- and middle-class people forced into regular use of dangerous vehicles by the demands of employment and daily living.

Similarly, a deeper consideration of medical negligence identifies the central role of medical entrepreneurs, corporate executives, lawyers and accountants of private medical and pharmaceutical corporations and their political representatives who have worked to prevent the development of comprehensive, efficient, accessible and accountable public health systems, ensured the growth of expensive

private health care, and continued to market dangerous or ineffective medical products on a vast scale.

The mass of working people, as principal victims of corporate power, currently sustain welfare provisions through income and consumption taxes. The wealthy minority who benefit from their suffering have ways to evade income taxation, including use of family trusts and political pressure for reduced taxation of unearned income. As executives they avoid corporate taxation through use of tax havens, transfer pricing and special concessions from sympathetic or supine governments.

At the same time, deregulation, international mobility and increasing concentration of capital create increasing pressure for wage reductions. This undermines the tax base of governments reliant upon taxing workers' wages and consumption, making it increasingly difficult to fund comprehensive welfare provisions. On the other hand, the scale of contemporary productive investment (and wealth generation) creates the potential for massive extension of general social welfare. We see here that the issue of no-fault compensation, or, more broadly, of decent social welfare provision, including a universal basic income for every citizen, is bound up with the issue of rational organisation of production and radical tax reform, forcing the wealthy to pay their share. This, in turn, is bound up with the need for policies aiming to reverse current trends towards corporate deregulation and unrestricted movement of capital across borders, privatisation of public services and corporate concentration.

Prevention

The greatest need is to take steps – beyond the limited deterrence effects of punitive sanctions and victims' compensation – to prevent pollution, environmental damage, accidents and injuries. As Abel points out, 'money cannot really restore victims to their status quo before the accident … reimbursement for the cost of medical treatment is hardly the same as never being injured … money is a poor equivalent for non-pecuniary loss'.[24]

Effective democratic regulation and control of production, with investment directed toward social welfare, rather than profit, could allow for improved conditions of health and safety at work, and a move to safer, more efficient public transport systems. Properly funded and regulated public health systems could reduce adverse

medical reactions. Reduction in the political influence of private corporations could allow for more effective legal protection of workers, consumers and the environment.

Epidemiological evidence of significant correlations between, for example, exposure to potential hazards and deleterious health or environmental effects should be deemed sufficient for immediate action to remove such hazards, prior to definitive evidence of a causal relationship. As suggested in chapter 5, lack of strong positive evidence of long-term safety or of social benefits exceeding costs should be treated as grounds for outlawing emission or sale of materials. The onus should be upon those who wish to release any new product or by-product or continue to release anything whose safety has been seriously questioned (on grounds of possible serious or irreversible damage to people or the environment) to provide such positive evidence or cease production and clean up, with strong criminal prosecution for unauthorised release.

As noted earlier, epidemiological evidence strongly suggests that those on lower incomes in societies with greater income disparities, particularly the unemployed, are more susceptible to accident, injury and illness than those in less unequal societies, as a result of stress-induced immune depletion, and loss of social cohesion, trust and collaboration.

Increased taxation of corporations and private wealth supporting expanded social welfare systems, along with government job creation, improved minimum wages and restricted executive salaries, could therefore be expected to significantly reduce accidents and injuries.

Granting all citizens sufficient income to satisfy basic subsistence needs would also radically alter the balance of bargaining power between labour and capital through offering ordinary people a genuine alternative to paid employment. It would thereby allow such people to demand decent treatment, respect, empowerment and healthy conditions of work.

Discussion topics

1. What are, or should be, the purposes of tort law?
2. Does the law of negligence achieve effective compensation and deterrence?
3. What are the problems with the new legislation?
4. What are the alternatives?

Additional resources

R.L Abel, 'A Critique of Torts', in H. Lunz and D. Hambly, *Torts, Cases and Commentary*, 4th edition, Butterworths, Sydney, 2002.

H. El Menyawi, 'Public Tort Liability: An Alternative to Tort Liability and No-fault Compensation' (2002) 9 *Murdoch University Electronic Journal of Law* 1–21.

Chapter 18

FREEDOM OF THE WILL AND CRIMINAL CULPABILITY*

Mens rea

In the common law, criminal culpability is understood primarily in terms of free individual decision and action. As Hart says, 'the principle of punishment should be restricted to those who have voluntarily broken the law'.[1]

From this perspective, the essential role of the criminal law is to protect the socially permissible free action of good citizens, centred upon 'enjoyment' of their life and property, through punitive redirection of the anti-social free choices of others. This emphasis upon free individual choice, as ground of culpability, finds expression in the requirement to establish that the accused's guilty mind, or mens rea, has caused a particular criminal action, or actus reus. Did they really intend to perform the action in question, and did their intention really cause them to perform the action?

The 'action' here refers to voluntary or intentional bodily movement – a physical movement that results from the operation of the will. Or, as Hart says, a bodily movement 'subordinated to the agent's conscious plan of action'.[2]

Mens rea has come to refer to a range of states of mind. But the underlying idea is that of a responsible agent who has chosen to break the rules. This includes a conscious – prior – decision to break the law, but can also include the intention to engage in action the subject knows involves a chance of causing a prohibited result, even though this is not the aim or purpose of the action.

Consciousness is therefore a necessary condition of criminal culpa-

bility. This is not merely consciousness in the broad sense in which animals, as well as humans, are 'aware' of their surroundings, and experience feelings of various kinds. Rather it is consciousness in the narrow – and possibly specifically human – sense of being able to stand back from our mental states and think about, appraise and evaluate them. As Richard Norman says:

> We need to do this in order to make rational decisions about our future states, by reviewing our various, perhaps conflicting desires, considering the reasons for and against acting on them and assessing the weight of these different reasons. [Because it allows us to plan and evaluate our actions, consciousness in this sense] is a precondition for our status as moral beings … Because we are conscious beings … we are not just prisoners of our immediate environment, but can [freely] choose our actions and thereby make them our own.[3]

The rationale of punishment as retribution is crucially dependent upon these ideas. The argument is that because the individual offender has freely chosen to break the law, so have they, therefore, in a sense, chosen to expose themselves to the possibility of state-inflicted pain and suffering as the cost of the suffering they have inflicted upon others. This idea of free decision is central to the rationale of punishment as specific and general deterrence. The infliction of suffering upon those found guilty of criminal offences aims to influence future free decisions by the individual concerned, and by others who might otherwise choose criminal, rather than legal, means, to encourage them to refrain from any such illegal choices.

Here we focus upon some major problems of the application of the idea of free individual decision-making within the criminal legal system, drawing particularly upon the work of the analytical philosopher John Searle. Searle's work is useful because of his clear picture of the nature of free will and free choice. But to understand this, it is necessary to first consider his analysis of certain aspects of the mind–body relation, centred upon the phenomenon of intentionality.

Intention

As Searle notes, intention in this sense refers to a mental state that is both a reason for action and a cause of action. Some intentions are formed prior to action, some are not. In the former case, we want to achieve X, we believe that we can achieve X by performing action Y. This desire and belief together provide at least part of the cause of our

deciding, and hence intending, to perform action Y. We have the intention to perform action Y prior to the performance of the action itself. But such a prior intention then becomes part of the cause of our performing action Y – along with access to relevant means and opportunity. The appearance of opportune circumstances – which trigger us into action – might be called the 'proximate' cause of the action.

When the action is performed, what makes it an action (rather than a mere movement) is the involvement of a mental component or intention-in-action, as well as a physical component of body movement. The prior intention causes the action, which itself involves the intention-in-action causing particular muscular contractions. Bill had a prior intention to kill his uncle by shooting him, he decided to do so, and this prior intention was part of the cause of the intention-in-action of his shooting his uncle, when the uncle stepped out of his house. His intention-in-action (to shoot the uncle) caused his finger to pull the trigger.

Often the prior intention will be of quite high order (getting to work) and require a complex sequence of specific subsidiary actions for its achievement (opening the car door). In fact, intentions come in nested hierarchies, with some higher-order 'plans' possibly taking years to unfold, through innumerable more specific sub-plans and intentional actions.

At the other end of the scale, there may be no prior intention at all. 'Many of the actions we perform, we perform quite spontaneously, in response to the circumstances in which we find ourselves, without forming any prior intention to do these things.'[4] But such spontaneous actions still involve the causation of bodily movement by intentions-in-action. They involve pre-existing beliefs that shape and direct our perception of the situations in question, and hence also our responses to such situations. And they still result in part from previous prior intentions and actions that have brought us into particular situations.

Free will

In Searle's model 'free will' enters the picture in the space between reason and decision. I want X and can see that doing Y will get it for me. But is Y really the right thing to do? Will it hurt others? Is it legal? This gap, as Searle calls it, between reason and decision allows me to consider other possibilities, obligations or commitments that might conflict with, or override, the performance of Y to get X. They might

more urgently command the resources necessary for accomplishing X – resources of time, effort, money or whatever. Y or X (means or end) might simply be morally or legally unacceptable.

This gap provides a space not just for adjudicating between conflicting reasons for action, but for rethinking the nature of our options, obligations and commitments in the light of new information and experience. Such rethinking need not just be a matter of internal thought processes. It can also involve discussion and debate with other people

Just as the gap between desire and decision allows for the possibility of deliberation, including reference to moral and legal considerations, prior to decision, so do the gap between decision and action and possible gaps within ongoing sequences of actions allow for further reconsideration. We can sometimes still 'change our mind' even after we have embarked upon a particular 'course of action'.

Such gaps do not always exist. We do not always form intentions to perform actions prior to the performance of the actions themselves. Spontaneous responses are partly a result of external situations that might not allow for deliberation (for example, being attacked in the street or suddenly confronting a changed traffic light in the path of our car), and partly a result of internal situations: our tendency or propensity for (particular) spontaneous responses in situations of the type in question.

Our responses are determined by our perceptions of situations, as situations of a particular type; a situation of traffic-light change, for example. And our perceptions are shaped by our established patterns of belief, including, in this case, our beliefs about the operation of traffic lights and the consequences of ignoring their changes of colour. Our perceptions and actions can be affected by preceding trauma, or consumption of psychoactive substances.

Earlier 'free choices' might be relevant in causing such spontaneous responses. If we know that we are likely to respond in particular ways in particular situations, we might choose either to retain such propensities or to try to change them. We might pursue anger management strategies to change undesirable responses of spontaneous violence or aim to improve our driving skills. We might choose to avoid the sorts of situations likely to trigger undesirable responses. We might choose to avoid consumption of alcohol or other drugs because of the likely or possible consequences for our perceptions and spontaneous responses.

Action and omission

Criminal law-makers in the common law system have been anxious to distinguish 'action' – as a source of moral and legal responsibility – from 'inaction' – which incurs no such responsibility. From this perspective, everyone is responsible for looking after themselves alone. No-one should be bound to be 'their brother's keeper'. Thus, unlike continental civil law, the common law recognises no general duty to rescue others in distress, even where this can be accomplished with ease and little threat to the rescuer.

Traditionally, this distinction has been defended by the assertion that it is the law's business to 'arrest acts of positive harm' but not to 'encourage – or require – acts of positive good'. It is supported by the right-liberal idea of society as no more than an aggregate of competing human individuals, any one of whom is responsible to any other only to the extent that they have voluntarily entered into legally recognised contractual relations with the individual in question.

A closely related issue is that of medical personnel causing death by 'failing to administer' necessary medication or nutrition, at the request of the patient or their representative, to terminally ill adults or seriously handicapped newborn babies. This is treated as lawful, in contrast to 'active' administration of a lethal injection at the patients' or their representatives' request, for example, which is treated as unlawful. Here again, the causal efficacy of 'inaction' is denied, with death attributed rather to the 'natural course of the illness'.

But it is far from clear that such ideas make sense in practice. It has frequently been pointed out that failure to take action is a form of action since the individual actively restrains themselves from taking the apparently 'correct' action, or actively withdraws treatment, in the latter sort of case. As Norrie says:

> an omission can be described as a negative act, a description that indicates that omissions are in their essence similar to rather than different from acts. Omissions can be conscious decisions either not to do something or to do something other than the thing that is not done … Either way, to describe a failure to act is as much to describe a practical orientation to the world as is the description of an act.[5]

If the real moral and legal issue is 'free decision' to do 'the right thing' there is no moral distinction between the two sorts of cases. Definite causal consequences, including 'positive harms' and 'goods',

flow from the decision not to take action, as much as from the decision to act. In Norrie's words:

> omissions can be as much the cause of an event as acts. An omission can serve just as well as an act as a necessary and sufficient condition for any particular outcome.[6]

This is particularly clear in the sort of medical case considered earlier. Here, abstaining from treatment has the same consequence as actively assisting suicide or engaging in voluntary euthanasia, in the form of the death of the patient. Furthermore, as acknowledged by the House of Lords in *Airdale NHS Trust v Bland*,[7] such legal abstention could also produce weeks of pain prior to death that could have been avoided by an illegal lethal injection.

In terms of Searle's analysis, 'inaction' in specific situations could be a result of a particular prior decision, producing a prior intention. We decide to refrain from action in some particular future situation, with a view to producing particular causal consequences of such inaction. Perhaps this involves 'actively restraining ourselves' from performing some action we might, otherwise, feel constrained to undertake.

Similarly, past decisions could contribute to 'spontaneous inaction', without any 'direct' prior intention. Failure to take steps to prepare ourselves to help others in need – as such occasions might arise – could contribute to our 'freezing up' when encountering situations in which we might otherwise have averted substantial harm to others, with minimum harm to ourselves, for example.

Reasons

As well as casting light upon the nature of 'freedom of the will', this analysis can illuminate the concept of 'reasons for actions'. The concept of a 'reason' implies some basis in rational deliberation. Prior intentions are products of deliberation concerning beliefs and desires. Such deliberation centres upon reasons why particular desires specify appropriate goals for action, why the individual concerned wants to do this, rather than something else; why this is seen as a priority, and/or a just and right thing to do on this occasion. It involves reasons why particular actions are seen as appropriate means for achieving the goals in question, why they are the best or only way to achieve such goals at the time in question.

Some goals will be ends in themselves; some will be means to other ends. When we ask for the 'reason' or 'motive' behind a particular action we are really asking what ultimate goal the individual was attempting to achieve, why they were seeking to achieve such a goal, and why they believed the chosen means to be appropriate to achieving this goal.

This is the subjective side of the reason for the action in question. But we cannot properly understand any such subjective considerations without reference to their objective circumstances and determinants: the particular objective social facts or conditions that have caused the relevant perceptions, beliefs and desires and facilitated or triggered the relevant reasoning processes. And this applies whether or not such perceptions, beliefs and desires are based in fact, or such reasoning is sound or cogent.

With spontaneous actions, the nature of the subjective response will depend upon the way in which the external situation is perceived. An individual's belief system will classify situations in particular ways, and this will determine that individual's likely responses in the situations in question.

Here again, we must consider also the objective circumstances of formation of such beliefs. Are such beliefs factually based? And if not, why not? Did the individual have access to reliable information, or have they been misled? We must consider the extent to which the individual concerned has been able to make rational judgments about the appropriate behaviour in the situations in question, or to act, effectively, upon such judgments. Did they have access to appropriate counselling or training for example? Or did they lack the resources to access such assistance?

Autonomy and the limits of free will

In the literature of ethics and philosophy of mind practical preconditions for, and restrictions upon, free will and free action are traditionally discussed in terms of 'autonomy' or self governance. Moral philosophers typically distinguish three different ideas or dimensions of autonomy that go along with different sorts (and degrees) of possible obstacles and restrictions to free, conscious action.

1 *Liberty of action*. An individual is autonomous in this sense if their action results from their own conscious intention and is not the result of external coercion or duress. When autonomy is identified

with liberty of action, the primary contrast drawn is between autonomy and coercion. Coercion involves the deliberate use of force or the threat of harm. The coercer's purpose is to get the person being coerced to do something he or she would not otherwise be willing to do.[8]

2 *Freedom of choice*. This refers to the range of real choices available to an individual, in terms of access to material means for the realisation of particular goals or desires. Thus, we noted earlier that intention was only ever a part of the causation of action. Such intentions only become 'operative' in effective causation of action if the individual concerned also has effective control of the necessary resources: of strength, skill, knowledge, tools, machines, assistance and opportunity.

3 *Effective deliberation*. This refers to internal, rather than external, resources available for, or restrictions upon, the exercise of individual autonomy, specifically to the individual's capacity for making rational and informed decisions. Such decision-making involves both rational choice of ends, in keeping with real need, for example, and rational choice of appropriate means for achieving independently chosen ends, in tune with what is logically, physically, and socially possible, and what is morally just and responsible.

> Powers of effective deliberation can be limited in many ways … some individuals may not have developed the necessary abilities or may even be incapable of sufficiently developing them … even individuals who have the requisite abilities may be unable to exercise them on a particular occasion due to various internal factors … emotions such as fear may make the impartial weighing of information impossible … the presence of pain or the use of drugs may also affect the exercise of reasoning abilities … Lies, deception, and a lack of appropriate information can all limit the effective exercise of the abilities required for rational deliberation.[9]

Defences

At first sight, this moral philosophical analysis of 'free will' seems quite compatible with the basic categories and procedures of contemporary criminal law.

First, the actus reus generally has to be an intentional action: a bodily movement that is the product of an intention-in-action.[10] The

law does not generally hold individuals responsible for bodily movements that are seen as caused by anything other than such an intention-in-action on the part of the individual concerned.

The legal category of 'automatism' refers to non-intentional bodily movements that occur when an individuals' conscious awareness is seriously impaired in some way, in other words, 'when the ordinary link between mind and body is absent'. Where the automatism is caused by something other than an 'unsound mind' or serious mental illness, such as sleepwalking, diabetes, major stress or a blow to the head, it can be seen to render action involuntary and hence not criminally culpable.[11]

However, as with the preceding analysis of spontaneous actions and unintended movements (as consequences of prior choices), those who freely choose to put themselves in situations where it is possible or likely they will be drawn into involuntary movements with criminal consequences – unintentionally discharging firearms during armed robberies, for example – can be held liable for such consequences. (See *Ryan v R* (1967) 121 CLR 205.)

Further, the criminal law does not hold individuals responsible for criminal acts that are seen as products of duress, in other words, where an individual commits an offence under threat of physical harm to themselves or another person. Here, the prior intention is, indeed, to perform the forbidden act – of theft or whatever. And this 'criminal' prior intention is the cause of a 'criminal' intention-in-action. But the goal or purpose of the action is to avoid serious (and unjustified) harm to the individual concerned or to another, rather than, for example, gain unjustified personal enrichment.[12]

Duress involves a severe restriction of the options available to the individual concerned. They are forced to make a 'coerced choice from morally unacceptable options'.[13] Where circumstances other than the threat of serious violence or harm by another person similarly restrict the options available to a person, such that they can only avoid 'irreparable evil' through the commission of some criminal act (involving a lesser evil), they can apply the defence of 'necessity' (or duress of circumstances).

Self-defence is understood in a similar way. The decision and prior intention to use force is justified if the individual believes, on reasonable grounds, that such force is necessary to defend themselves against an unwelcome attack.

Where the individual's internal decision-making capacities have

been radically compromised by a mental illness or by some 'abnormality of mind', the law allows defences of insanity and diminished responsibility or substantial impairment The former allows for excusing the individual of criminal liability in any offence. One or the other of latter is available in some jurisdictions in relation to murder.[14]

The defence of provocation relates to spontaneous action, triggered by acts or words of the victim that produced 'a sudden and temporary loss of self control'. Here the law excuses spontaneous responses that it is unreasonable to expect the individual to take prior action to avoid.

The criminal law's treatment of intoxication also appears to be in tune with the preceding analysis of individual responsibility in relation to spontaneous actions. Individuals are not criminally liable for actions committed in extreme states of intoxication – where their action ceases to be voluntary – if such intoxication does not result from their own free choice. If they have 'freely chosen' to become intoxicated, then they have chosen what they know to be possible criminal consequences.[15]

A black-and-white picture

But while the law is in line with the broad categories of the moral philosophical analysis, it departs from such analysis in its crudely black-and-white approach. Actions are regarded either as completely free – and hence culpable – or completely (or largely) determined by internal and/or external circumstances, and hence not culpable.

This is particularly evident in relation to insanity. Ideas of legally relevant insanity have been restricted to extreme states of mental illness where the defendant was, at the time of the crime, 'labouring under such defect of reason, from disease of the mind, as not to know the nature and quality of the act he was doing, or if he did know it, then he did not know that what he was doing was wrong'.[16] In effect, this test only covers those defendants who, in extreme cases of mental illness, are unaware of what they are doing or of the significance of their actions.'[17]

We might think that diminished responsibility – or substantial impairment – would cover lesser mental problems. But this defence is available only in relation to murder charges. It is not available at all in some jurisdictions. It merely reduces murder to manslaughter. And juries can return a verdict of murder even if there is medical evidence

of the appropriate sort of mental abnormality or impairment.

In effect, the only sort of duress by the action of another, recognised by law, is 'the threat of immediate death or serious personal violence', where such immediate threat either 'overbears the ordinary power of human resistance' or provokes legitimate resistance in self-defence. The only sort of duress of circumstances – or of 'necessity' – recognised in law is a threat of immediate death or very serious injury by such natural disasters as fire or earthquake.

Provocation is a defence only in relation to murder, and only applies to actions of the deceased, immediately prior to the killing, 'of such a character as to cause an ordinary person to lose his self-control to such extent as to act as the killer has acted'. And though the NSW *Crimes Act* has recently been updated (in s 23) to include provocative actions of the deceased 'at any time before' the homicide, this is intended only to accommodate histories of ongoing domestic violence in relation to battered woman syndrome.

Beyond this point, the only idea of action by others 'pushing' individuals into criminal behaviour is that of unintentional and extreme intoxication by another's action.

But it is clear that, in reality, there are many different forms and degrees of freedom and determinism, autonomy and restriction, many sorts of ways in which the actions of some affect and constrain the decisions and spontaneous actions of others, increasing the likelihood of criminal behaviour. Not only does the black-and-white legal picture provide a false dichotomy, it also presents extremes that actually make little sense. There is no absolute and unconditional freedom, nor any pure determinism where human actions are concerned.

Motives

Most important, we cannot make sense of specific cases of choice as free choice without reference to the detailed processes of decision-making involved. This is, after all, the essence of 'free will', on the analysis developed so far. Free will exists where an individual is in a position to make a real choice between genuine alternatives on the basis of rational – and informed – deliberation. Such deliberation generates the reasons for their action. And only by examining the details and context of specific decision-making can we rationally access the extent and nature of restriction of the processes in question. Yet the criminal law steadfastly refuses to properly consider any such decision procedures.

As Norrie says:

> It is as firmly established in legal doctrine as any rule can be that motive is irrelevant to responsibility; a crime may be committed from the best of motives and yet remain a crime.[18]

'Motives' typically enter the discussion only at the level of prior intention, where such intention is part of the definition of the offence. Did the defendant really intend to permanently deprive the victim of their property? Was this their motive or reason for taking it? In other words, there is no consideration of the ultimate goal of the action, of the reasons why this goal was chosen or why these means were employed.

Sometimes, what Norrie has called a 'utilitarian psychology' of motive is seen as useful in establishing the relevant facts of the case. But this sort of analysis is very different from the earlier consideration of motive and reason, with its emphasis upon rational deliberation, belief and context. In this system of thought

> each individual is seen as a separate monad operating according to discrete personal motivating characteristics or emotions … Individual psychic forces like 'jealousy' or 'greed' or 'anger' or 'love' are seen as the 'springs of action'. No thought is given to the social context within which 'jealousy' or 'greed' are stimulated or to the particular content they embody.[19]

As Norrie concludes, legal doctrine has become so structured that

> intention becomes crucial to liability and motive becomes irrelevant. Then, at the end of the trial, when justice has been done, and a criminal has been properly convicted, the doctrine can be put to one side and motive allowed back into the courtroom, in the non-threatening guise of a factor in mitigation of sentence.[20]

The historical basis for this development is clear:

> Desperate social need and indignant claim to right were the motives of the poor in the seventeenth, eighteenth and early nineteenth centuries. These were hardly motives calculated to win favour or compassion from a social class determined to impose a property order on all regardless of the consequences.[21]

The only situations where the defendant's reasoning process and resulting goal or purpose (in performing the criminal action) is considered relevant to the issue of culpability are those involving duress and self-defence. As Mousourakis says, such cases involve 'a morally

worthy motive; that is, self preservation'.[22] But these ideas are applied in such a narrow and restricted fashion as to be irrelevant in the great majority of criminal prosecutions.

In reality, reference to jealousy and greed, anger and love, 'desire for money and perverted lust', contributes little in the way of real explanation of criminal activity. Everyone experiences such thoughts and feelings more or less frequently. The real issue, as Norrie points out, is how and why such feelings come to take the particular forms they do and contribute to particular consequences, criminal or otherwise. And to understand this, we must consider the social context and psychological makeup of the individuals and the actions concerned.

Sociological issues

On the sociological side, we need focus only upon the broad reality of contemporary economic life. In a capitalist firm, the owners have ultimate authority. They are interested in profits and in maintaining the conditions that keep the profits rolling in. They cannot oversee every aspect of workplace activity to ensure that these conditions prevail so they hire special employees who are empowered to act on their behalf in the day-to-day organisation of such profit maximisation. Such co-ordinators oversee the day-to-day actions of the workers – the actual producers and wealth generators – whose jobs are, in consequence, hugely restricted, controlled and lacking in creative possibilities. They obey the co-ordinators out of fear of being punished or fired.

A hierarchical division of labour within larger and more complex corporate structures apportions tasks, empowerment, status, remuneration and quality of life in hierarchical order. A few at the top have excellent working conditions and substantial empowerment. They are autonomous in that they are largely free of day-to-day external duress, they have freedom of choice – in terms of real options in the organisation of their work and leisure time – and they have access to a wholly different quality and quantity of information for rational decision-making within the organisation, including access to substantial human and mechanical information processing resources.

The economic power of owners and higher co-ordinators gives them power to directly influence the political process, participating in the formulation of legislation and directing the day-to-day decision-making of the political leadership. They can exert huge social and political power via the control of mass communications.

Those who have acquired sufficient assets through inheritance or other means can live very comfortably with no need to work at all. They can leave the management of such assets to others. They can choose to do whatever work might appeal to them, or simply enjoy the benefits of consumption. They can study what they want, possess what they want, travel when and where they want.

Those in the middle ranks fall well below these levels of autonomy, in relation to both effective power and decision-making within the organisation and the power of money outside it. And the great majority, in the lower ranks, have effectively no power at all, within the structure or outside it. In the first world they probably do have access to basic necessities of food, clothing, accommodation and transport, though the increasing numbers employed on a casual or part-time basis find it more and more difficult to make ends meet.

Many of those at the bottom do not enjoy their jobs, or do not enjoy much of what they do at work. They work in order to get money to live. While they work they follow orders from others, with which they do not necessarily agree. Often they have no knowledge of the decisions underlying such orders, who has made them, how or why. And today, with high levels of chronic unemployment and underemployment, they must struggle all the harder – through appropriate deference and ever more unpaid overtime – in order to hold onto any kind of employment. In other words, the greater part of their working lives consists of coerced, rather than free, action. They operate under continuous duress. And their limited incomes similarly reduce the options available to them outside the working environment.

Beyond the corporate sector is the small business sector where owner-managers exercise some power over their day-to-day lives but struggle to keep their heads above water in an intensely competitive environment. They have to pay the monopoly prices of big corporate suppliers, and struggle to compete both with other small businesses and with the cartels, economies of scale and long-range planning of big corporate operators. And conditions for workers within such smaller enterprises are typically even worse than those in bigger operations because of much tighter profit margins, higher costs, negligible unionisation and weaker legal protections.

Beyond the small business sector is the world of the long-term unemployed and radically underemployed. Government social security payments fail to accommodate rents, food, clothes, etc to allow for even minimally decent living standards. Some do not even receive

these payments and some are unable to make best use of them. In other words, their life options are severely restricted – so long as they stay within the legal economy. And powerful social forces of denigration and victimisation, by the police and social welfare authorities amongst others, work to undermine the self-esteem and psychological wellbeing of those in this group.

Advertisers and marketers make massive efforts to stimulate ever more wants for ever more consumer goods. New products can be presented as compensations for the general powerlessness, frustration, insecurity, domination and struggle of the majority of the population. But such potential consumers feel all the more inadequate and frustrated when they cannot – legally – access such proliferating 'compensations'. Here is another major source of duress for many people – the duress of thwarted desires, along with envy and resentment of those who apparently have everything while doing nothing to deserve it.

Supposedly, the advanced industrial democracies allow and encourage 'mobility', whereby, through appropriate effort, those born into the lower echelons of the social class structure can drag themselves up to higher levels. Anti-discrimination laws are supposed to safeguard and facilitate such mobility.

In reality, available positions rapidly decline up the hierarchy. No matter how much effort is expended by the lower orders, only a tiny percentage will be able to ascend and only at the expense of others moving in the opposite direction. Anti-discrimination legislation cannot compensate for the advantages of established wealth, influence and connections enjoyed by those born into the upper echelons.[23] Some individuals, through luck, as much as effort, through appeasing and pleasing those above them, are able to ascend a step up from the lower levels. Some choose not to try to do so, given the behaviour required of those in higher levels. Most remain in the social stratum into which they are born.

Class and crime

The hierarchy of the social class structure therefore corresponds to a hierarchy of degrees of liberty of action, freedom of choice and effective deliberation, with reduced free will and increased duress further down the system. And such a structure goes a long way towards explaining patterns of crime in contemporary society.

At the top, those who operate and benefit from such a fundamen-

tally unequal and immoral system have the capacity to commit crime on a grand scale with effective impunity. The very scope of the freedom enjoyed by members of the ruling elite goes a long way toward explaining their proclivities for price fixing, collusive tendering, bribery, patent violation, false advertising, insider trading, trading while insolvent, paying below-award wages, tax and other fraud, manufacture and sale of dangerous products, environmental damage and provision of unsafe working conditions leading to industrial death, injury and disease. In their role as agents of the duress experienced by those below, they should also be held responsible as inciters of, and accessories to, the crimes of those below them.

In the lower ranks, it is restriction of the scope of free choice and action that explains predominant patterns of crime. The intense competitive pressures upon smaller businesses contribute to crimes in the interests of cutting costs, including employment offences, consumer offences, food offences, and environmental crime. 'Poor economic performance [can] make offending seem a rational choice to maximise profits or to ensure the survival of the business.'[24]

Within the working class a proportion of those who are continually reminded that they are without value in the legal economy may turn instead to the criminal economy. A proportion of those who are continuously victimised by the forces of law and order eventually lose respect for such a system, and become involved in street crime.

Some of the most frustrated, exploited and downtrodden members of society may turn to illegal drugs for solace, as well as to alcohol, cigarettes and legal anti-depressants and stimulants. Those who become addicted to such illegal drugs but cannot afford to pay for them through legal employment turn to crime to support their habits. Addiction is a potent form of duress, analogous to immediate threats of violence.

For some of those faced with the poverty and ignominy of long-term unemployment or a life of powerless drudgery with minimum respect and remuneration, a life of property crime or drug dealing can appear to offer a viable alternative, with mental challenge, excitement, financial reward, independence, empowerment and self-respect.

The two most common types of interaction identified as crimes of violence by the Australian criminal justice system are 'confrontational violence between males, typically young and of marginal socio-economic status', and 'violent interaction between family members and other intimates', both often involving alcohol.[25]

In the first case, we can trace a path of causal determination from income inequality and discrimination, through disrespect and powerlessness, to street violence associated with the defence of honour. The greater the scale of income and social power inequality, the more those at the bottom of the scale, experiencing comparatively greater poverty and powerlessness, and corresponding social disrespect, may feel that they have to defend the vestiges of self-respect they have left. Physical violence is sometimes seen as the only means at their disposal to do so. In the second case, as Hogg and Brown observe:

> it would be surprising if the material stresses generated by poverty, unemployment and social isolation did not seriously exacerbate the ordinary, day-to-day tensions that arise within family relationships and produce higher levels of conflict and violence.[26]

At the same time, it needs to be borne in mind that, as Wilkinson says, violence will always appear to

> [be] concentrated in poor areas and occur primarily amongst the poor themselves … because what counts as violence are those forms of coercion not sanctioned by social institutions; making people homeless by ending a tenancy is not an act of violence, whereas hitting the landlord is.[27]

The empirical evidence supports the idea that duress of social deprivation motivates a significant proportion of street crime. A prison survey in 1973 found that

> only 4% of prisoners had completed the HSC, compared to over 20% of the general male population aged 15 years or more. More than two-thirds belonged to the lowest occupational strata [unskilled], compared to just over two-fifths of the general adult population; a further 30% belonged to the second lowest stratum [clerical, trades skilled] and under 2% were from professional or middle management backgrounds compared to almost one-quarter of the general population.[28]

And a survey in 1996 found that

> 50% of male prisoners and more than 75% of females had not been employed in the six months prior to their imprisonment. More than 50% of males and almost 50% of females had not completed secondary schooling to school certificate level and fewer than 10% had experienced any post secondary education.[29]

Indigenous people are heavily overrepresented in the social underclass, with unemployment levels as much as 80–90% in some communities. As Hogg and Brown note:

> When at work, Aboriginal people earn on average about half of the income of non-Aboriginal Australians … government payments [are] the main source of income for [over 50%] of Aboriginal people … They are massively disproportionately over-represented amongst the homeless population and the educationally disadvantaged. Aboriginal communities continue to experience high levels of [substance abuse,] infant mortality and health problems.[30]

And, 'Indigenous people in Australia are incarcerated at massively disproportionate rates compared to the general population.'[31]

> As at the 2000 prison census, there were 4095 Indigenous prisoners in Australian prisons, constituting 19% of the prison population. The national rate of Indigenous imprisonment was … almost 15 times the non-Indigenous imprisonment rate.[32]

Psychological issues

On the psychological side, individual life paths, including free decision-making, are profoundly shaped by personality structures established in early childhood. And different sorts of personalities impose different sorts of restrictions upon individual autonomy and free will.

Particularly significant are issues of self-esteem. Those who do not achieve a 'built-in sense of self-esteem' in childhood tend to develop depressive personality structures, leaving them vulnerable to self-blame and feelings of hopelessness and worthlessness in face of reverses and difficulties. Such individuals need repeated assurance in the form of others' good opinion in order to maintain their own psychic health.[33]

Given the potentially disastrous personal consequences of the negative judgments of peers and others for those prone to depression, such a need for recognition and reassurance can override both principled objections to criminal activity and fear of criminal penalty in motivating such individuals to participate in joint criminal operations, doing their part to retain the love or respect of other members of the gang or criminal 'community'.

Because depressive personalities tend to suppress their own opinions and defer to others in their attempts to win love and approval, they tend to accumulate an increasing amount of (repressed) resentment, anger and aggression. There is always the danger that such aggression will be turned away from themselves and onto others, particularly those they feel have let them down, or those who they feel they can safely abuse.

Hysterical personalities are similarly 'dominated by the urgent need to please others in order to master the fear of being unable to do so'. As a result of being disregarded in childhood, their needs not appreciated, they become demanding and attention-seeking as adults. 'This results in restless activity, dramatisation and exaggeration … And unrealistic dependence upon others.'[34] 'People with hysterical personalities have high anxiety, high intensity and high reactivity.'[35] Their proneness to drama and risk, and to trying to master frightening possibilities by initiating them, can get them into criminal legal difficulties.

At the other end of the scale are those obsessional personalities driven to pursue their own priorities – of order, organisation and control – irrespective of what others might think, and sometimes at others' expense. Driven by their own conscience, rather than by the opinions of other people, it is quite possible for such individuals to fall foul of the criminal justice system, where its values do not happen to correspond with their own.

For some, the voice of conscience is the voice of God, commanding absolute obedience, even if this is at the expense of being seen – by some – as a criminal and a law-breaker. Direct sanction by divine authority can sometimes be seen to justify or demand horrific acts of violence.

A powerful, rigid and punitive conscience can be firmly established in early childhood on the model of parental authority, rewarding the individual concerned with feelings of pride for conforming to its dictates and punishing failure to do so with pangs of guilt. Its value system can be out of kilter with the more developed ideas and values of the adult. In such a situation the conscience itself is experienced as exerting powerful internal duress upon the individual concerned, forcing them to engage in – or avoid – particular actions 'against their better judgment'.

Schizoid personalities are still further removed from influence by group norms and values. By virtue of having lost a care-giver in early life, or having been treated as an appendage to the parents, or as someone whose needs are destructive of the person to whom they turn to fulfil them, they have withdrawn from intimate relations with other people (as too painful, dangerous or overwhelming). They have developed a complex fantasy world to compensate for their lack of fulfilment in the real world. And to the extent that fantasies of power and domination, including paranoid feelings and delusions, spill over into their actual relations with other people, such individuals can come into conflict with the criminal law.

Particularly serious problems of early childhood development and/or later social experiences lead to more extreme forms of such behavioural tendencies and greater likelihood of more serious psychopathology. And such psychopathology runs along a continuum from mild and temporary neurosis to serious long-term psychosis.[36]

Tangential to, but interacting with, such issues of personality are issues of intellectual development and disability. Here again, some individuals will be significantly disadvantaged by their heredity, or by developmental damage, and will not necessarily receive the special assistance they might need to realise their full potential.

Again, the empirical evidence confirms the significant involvement of psychopathology and intellectual disability in the causation of crime. The 1973 prison survey, referred to earlier, found that one-third of prisoners had received professional help or treatment for a nervous, emotional or mental problem. Nor does this mean that other prisoners were necessarily free of such problems. And studies by Hayes and McIlwain in 1988 and 1997 found 12.9% of the NSW prison population as having an intellectual disability.[37] Levy observes that

> the prevalence of mental illness among prisoners is higher than in the general community. In a submission to a recent NSW Parliamentary Select Committee Inquiry on the Increase in Prisoner Population, the Department of Corrective Services stated that 13% of female inmates in New South Wales have an intellectual disability, 21% had attempted suicide and 40% had a diagnosis of personality disorder.[38]

Interactions

Such psychological and sociological processes of disadvantage are intimately interconnected. Even relatively mild dispositions towards depression, compulsion, schizoid withdrawal or paranoia can lead to serious pathology under the pressure of social circumstances of discrimination, deprivation and powerlessness. And those suffering such deprivation are least likely to get effective assistance in coping with such pathology, or with intellectual disability.

The stresses of social powerlessness and anxiety in the parental generation can create major psychological problems for new generations, insofar as the latter become victims of parental mortality and morbidity, psychopathology, drug abuse, violence, disintegrating family structures, poor housing and nutrition.

Teenagers of all classes are driven by hormones and social dislocation into risk and challenge to established authority. But social and

psychological disadvantage brings working-class youth into contact with the criminal justice system at an early age, whereas middle- and upper-class youth are protected. Significant numbers of such disadvantaged young people move from juvenile detention centres on to adult prisons. The appalling conditions of the latter, with inmates stripped of all human rights and human dignity, subjected to regular threats and assaults by staff and other inmates, exposed to dirty air, poor-quality food and rampant hepatitis C infection, with huge pressures towards hard drug use in dangerous, insanitary conditions, exacerbate social and psychological disadvantage. [39]

Inmates often emerge psychologically and physically damaged, with criminal reputations, which make them the objects of intensified police surveillance, harassment and victimisation, and make rewarding employment – or any legal employment – all the more difficult to obtain. Hence the high recidivism rates, as victims become trapped in a cycle of ever diminishing real options and further reduced 'free will' and autonomy, with 56 per cent of prisoners known to have served a prison sentence on a previous occasion on prison census night in 2000.[40]

At the other end of the social scale, the wealthy and the powerful have access to the best resources for coping with both mental illness and intellectual disability, including long-term humanistic and insight-based therapies, rather than debilitating and toxic drug treatments. At the same time, they are free to indulge the more anti-social tendencies of their psychopathologies without fear of criminal legal intervention.

Wealthy narcissists can command the uncritical approval and respect – indeed love and worship – of others, which they need to maintain self-esteem, or, at least, they can command the appearance of such love and respect, through appropriate positive and negative reinforcement procedures applied to those within their power.

Wealthy obsessives and compulsives can indulge their will to power, domination, control, discipline, order, and cleanliness through rigidly disciplining and controlling the lives of those subject to their authority, including workers driven to ever greater productive efforts, of harder work and longer hours.

Wealthy paranoids can indulge their need to oppose the forces of evil through seeking, and achieving, political office, or power within law-enforcement agencies. And serious psychopathology and intellectual disability seem to pose few problems for successful careers at this level of the social hierarchy.

Over a dozen research projects on corporate career mobility

demonstrate that psychopathic or anti-social personality is an advantage when it comes to promotion to the 'top positions' of the corporate hierarchy. As Box points out, such research shows that those who attain such positions are

> not so much intelligent as shrewd – their organisational sense enabled them to sniff out the golden chance and grasp it firmly … they had the moral flexibility to meet shifting organisational demands and still enjoy the sleep of the just – their ability to relativise other moral imperatives while constantly prioritising the pursuit of organisational goals … [facilitated] a moral flexibility others denied themselves.[41]

Croall notes that 'personality types … associated with business success – including a propensity to take risks, recklessness, ambitiousness, drive, egocentricity and a hunger for power … are [also] linked with white collar crime'. Such crime can be justified as 'being in the interests of the company'.[42]

Criminal minds versus class structures

In this context, we can see how generally meaningless and misleading is any idea of the 'criminal mind' as 'cause' of crime. Were those at the top to be exposed to the circumstances of those at the bottom, no doubt they would respond in similar fashion. Their minds are no different; it is their social circumstances that are different, leading to different behaviour or different treatment for the same behaviour.

Supposedly everyone is equally subject to the same laws in all common law jurisdictions. Judges and magistrates swear oaths to 'do right by all manner of people … without fear or favour, affection or ill-will'.[43] And legal authorities have typically seen equal access to justice, including access to appropriate professional assistance, as an integral part of the 'rule of law'.

One might think that such equality implied treating those in similar situations in similar fashion. And, on the face of it, this is what the criminal law does. Supposedly, individuals are punished in proportion to the seriousness of their crime.

But 'seriousness' is ambiguous. It can mean the extent of actual harmful consequences. It can refer to something like the objective extent of risk of harm to which others are exposed by the action – how likely it is that such harm will eventuate and how great the possible harm is. Or it can refer to the extent of free will and autonomy – including the extent of restriction of such free will, on the part of indi-

vidual perpetrators – along the lines considered so far. Extent of actual harmful consequences creates major problems in a class divided society. Stealing from the rich can cause far less suffering than stealing similar amounts or much lesser amounts from the poor.

At the moment, for the crimes of the working class, ideas of objective harmfulness are built into the definitions of the offences – as murder or manslaughter or theft or burglary or robbery – without any reference to such social contextual considerations. So those found guilty are supposedly punished for the extent of real harm caused, without any analysis of real harm, at least until sentencing.

Their employers, on the other hand, in the crucial areas of health and safety at work and pollution, are held responsible for the supposed extent of risk to which they have exposed others, through contravention of relevant regulations, without reference to actual harmful consequences up until the sentencing stage. With risk treated as inherently less 'serious' than actual harm, this de-emphasises the criminality of employers as against workers. Together, these operations effectively preclude any real utilitarian assessment of the extent of harm caused by either the rulers or the ruled – within the criminal legal process.

Such comparatively 'lenient' treatment of the crimes of the powerful can be seen as offset by strict and absolute liability, with no requirement to establish a subjective fault element, while such a subjective element does have to be proved in working-class property crimes and crimes of violence. This only serves to mask the extent of culpability of the employers, in terms of their generally greater scope of free action, compared to that of the workers. At the same time, the limited application of available defences, with no real consideration of the duress of social deprivation, radically restricts what actually has to be proved in this area of culpability for working-class crime.

If we really believe in punishment in proportion to culpability, where the latter is a function, not only of the seriousness of the crime, but also of the degree of real freedom of choice able to be exercised by the perpetrator, then we must take account of the disparities of autonomy and freedom of choice across the social class structure.

Action motivated by need or desperation, by enforced ignorance or lack of available alternatives, is different from action motivated by pursuit of wealth and power by the already wealthy and powerful, and chosen despite awareness of both its harmful consequences and of real non-criminal alternatives for achieving similar ends. Action, which is not only criminal but also betrays the special trust associated

with social positions of power and privilege, would seem to carry greater moral culpability.

Beyond this point, it is unequal to impose punishments determined either by the seriousness of the crime or the extent of fault of the perpetrator, without reference to the circumstances of the individual punished. A fine can have quite different practical significance when imposed upon a rich person as against a poor person. A prison can have quite different consequences when imposed upon a vulnerable and attractive young man, likely to suffer rape in prison, or a hardened criminal psychopath likely to be the agent of such a prison rape.

Mandatory sentencing policies highlight, concretise and exacerbate the irrationality and immorality of the system. But past practices of judicial discretion in sentencing show little evidence of seriously addressing the issues raised here. Nor should the issues be left to the whims of such discretion.

As is suggested by the foregoing analysis, effective reduction of street crime (and, for that matter, corporate crime also) depends upon major social-structural change, addressing the inequalities of wealth and power that are the primary causes of such crime.

Discussion topics

1. Explain the role of free will in criminal liability. Include reference to omissions.
2. What is meant by a 'black-and-white picture' of free will and determinism?
3. How are issues of free will associated with issues of social class and or psychology?
4. Is there one law for the rich and another for the poor?

Additional resources

S. Box, *Crime, Power and Mystification*, Tavistock, London, 1983.
D. Brown and M. Wilkie (eds), *Prisoners As Citizens*, Federation Press, Sydney, 2002.
H. Croall, *Understanding White Collar Crime*, Open University Press, Buckingham, 2001.
R. Hogg and D. Brown, *Rethinking Law and Order*, Pluto Press, Sydney, 1998.
A. Norrie, *Crime, Reason and History*, Butterworths, London, 2001.

Chapter 19

CRIME AND PUNISHMENT

In this chapter we focus upon some of the ways in which classical liberal and libertarian thinking have most profoundly influenced the common law criminal justice system. Specifically, we consider classical liberal and libertarian ideas of legitimate punishment for criminal wrongdoing, based upon ideas of 'just deserts'. Just as in the economic sphere, libertarians have traditionally favoured the idea of reward in proportion to valuable social contribution, so, in the sphere of law, have they favoured the idea of punishment proportional to the extent of harm inflicted.

As we have seen in earlier chapters, in the tradition of classical liberal and libertarian thinking, punitive state intervention is justified only in the case of violation of others' core 'natural' rights to life and property. This means that for consistent libertarians, victimless crimes should not be treated as crimes at all. On the other hand, we have also seen how libertarians subscribe to a generally negative and pessimistic view of human nature, as greedy and self-serving, implying that such violations are likely to be fairly frequent without sufficient deterrent measures put in place by the state.

At the same time, such selfish nature is (generally) not seen to override individual 'freedom of choice' of action as cause of criminal wrongdoing. Rather, such violations of basic rights are seen as (typically) produced by the free choice of individual violators, involving some sort of decision about the costs and benefits of such actions to themselves. Libertarians require that such individual rights violators be held fully responsible for their criminal choices, without any refer-

ence to social or biological determination of their actions. And all such wrongdoing should be treated in a consistent fashion, so as to allow rational decision-making, and effectively influence such individual cost benefit analysis, in the direction of crime reduction in the future.

First and foremost such responsibility and consistency are understood in terms of *retribution*, in the form of infliction of suffering commensurate with the seriousness of the wrong the perpetrator has inflicted upon others, along with possible compensation paid by the offenders to their victims. Such state-imposed suffering takes the form of deprivation of property (fines), of liberty (imprisonment) and, in some jurisdictions, of life (execution).

Such privation and suffering, seen as consequences of transgression, are also supposed to serve the function of both *individual and general deterrence*, dissuading the individual concerned from committing further offences (on the basis of a revised calculation of costs and benefits), and similarly dissuading others from committing any such offence for fear of similar punishment. Imprisonment can also serve to protect life and property through *incapacitation*, depriving the imprisoned individual of the opportunity for further crimes (in the wider community) through confining them to the precincts of the prison.

As White and Perrone point out, in line with general libertarian concerns to minimise state power:

> this approach also supports the idea that security, law enforcement and prisons should be private rather than public institutions. This reflects a broad ideological commitment to the so-called free market as the best and most efficient avenue for the provision of social services.[1]

More efficient private services will cost the public less in taxes. Ideally, ways should be found to enable such systems to pay for themselves, through the productive labour of the prisoners or through charging them for prison services.

Historical considerations

The legal system in the English-speaking world has been profoundly influenced by classical liberal/libertarian ideas since the first development of such ideas in the 17th century. As Norrie notes:

> the revolution against absolutism came early in England, and the common law was instrumental in enshrining the rights and liberties of

> the subject against the sovereign and of establishing the role of judges as impartial interpreters of the law.[2]

In the 17th and 18th centuries, however, particularly draconian criminal penalties were imposed for any and every infringement upon the private property rights of the wealthy, with little consideration for the rights of offenders, or the justice or fairness of the punishments. Public displays of torture and execution were 'designed to manage the population at large, through fear and threat'. As White and Perrone observe, 'this approach to punishment [was] particularly concerned with the broader deterrent effect of exemplary sanctions, regardless of whether or not these [were] proportional to the crime'.[3]

But at the end of the 18th and in the course of the 19th centuries, under pressure from liberal, and particularly utilitarian, reformers, the focus increasingly shifted to the idea of 'doing justice to individuals' accused of law-breaking as well as to the victims or potential victims of crime.

This meant restricting culpability and punishment to those who had voluntarily broken the law, in the sense of freely choosing to do so, and consciously intending the criminal action in question, with knowledge of the likely harmful consequences of the action. In theory at least it meant applying the same principles, equally, to all offenders. And it meant applying punishments appropriate to the nature and extent of the harm caused.

Utilitarians, in particular, argued that only pain just sufficient to achieve effective deterrence should be inflicted, and no more. 'All punishments were pains and therefore evils in themselves, but they were justified because of their potential to deter the pain and evil of crime.'[4] As Norrie says:

> Translating [the classical] economic theory [of the free and rational utility maximiser] into the realm of social control, they argued that free individuals could work out for themselves that the costs of punishment must outweigh the benefits of crime and would, therefore, rationally desist from its commission … Just as the market regulates individual economic actions, so the criminal law regulates social control as an adjunct to the market, The task of law, according to Bentham, was to secure the harmony of individual interests, to supplement the invisible hand of the market, by keeping egoism within acceptable bounds.[5]

It is easy to see problems here, not least the gap between the real social circumstances of those likely to be accused of criminal offences at the

time – 'living in conditions that encouraged little respect for the social order and its laws'[6] – and the model of the 'freely choosing', rational calculator at the foundation of these ideas. Bentham himself recognised that, in fact, it is the conditions in which people live that determine their actions, particularly conditions of poverty, inequality and desperation, which determine the working class to do what they know is illegal despite harsh punishments (for example, stealing to live), so that crime reduction requires also substantial social-structural interventions by state power.

There are obvious parallels here with the problem of determining individual contribution to the value of productive output in the context of corporate structures and infrastructures, as the supposed basis for legitimate reward in the economic sphere, according to classical liberal ideas. Just as a complex constellation of interacting individuals is typically involved in the production of any particular finished good in an advanced industrial society, so is a similarly complex constellation of people involved in producing the harms classified as crimes.

As Box argues:

> conventionally defined crime is likely to rise amongst the poorest and most deprived sections of society, particularly in a time of recession. What is significant amongst those who are likely to turn to crime is the coupling of 'thwarted ambition' with 'relative deprivation' in a situation of marginalisation from institutionalised organisations of social change, and alienation from the forces of law and order.[7]

Therefore, all those who intentionally, recklessly or negligently contribute to such conditions of poverty, deprivation and alienation are, in part, responsible for the resulting criminal activity. This is likely to include politicians (particularly neo-liberal and libertarian politicians) and executives whose planning policies have directly contributed to increased unemployment, reduced access to social services, greater income inequality, and reduced social cohesion, trust and respect (including encouragement of racism and prejudice). It is likely to include police officers whose victimisation and persecution of particular social groups, such as working-class youth and ethnic and cultural minorities, has contributed to the alienation of such groups from established principles of law and order.

It is likely to include judges and prosecutors who send such victims of police discrimination to prisons where they become victims of further violence and abuse and acquire criminal reputations, connec-

tions and habits that increase the likelihood of further criminal involvements. And it includes those advertisers and marketers who contribute to the production of unlimited desires for expensive consumption items amongst those who cannot possibly legally acquire the means to purchase such items.

No doubt such individuals will respond by arguing that their contribution to crime was an unfortunate by-product of other useful activities whose social value outweighs the social costs of the crimes in question. Quite apart from the likely difficulties of establishing any such thing, libertarians have, anyway, strongly supported a general exclusion of considerations of motive from determination of criminal culpability. At best, such issues are to be addressed only at the sentencing stage.

There are also problems of appropriate matching of retributive punishment and crime, once we move beyond the old *lex talionis*, an eye for an eye and a tooth for a tooth. How do we compare the pain of losing a certain sum of money, for example, with that of a certain period of imprisonment? And from a utilitarian point of view, what if the extent of pain that needs to be inflicted to deter a particular sort of criminal activity is much greater than the pain inflicted by such an act? Presumably, no such penalty can be justified in such circumstances. It is fair enough to say that no-one should benefit at someone else's expense. But how much pain is permitted to avoid this?

There are issues here of how the extent of harm or pain to the victim is measured. Clearly, a theft of the same (small) sum of money from a very wealthy person will cause a lot less (objective) harm or (subjective) pain to that rich person than the theft of the same sum from a poor person will cause to that poor person. Indeed, stealing from the rich person to give to the poor person could cause very much more good (and happiness) than bad (and pain). Yet libertarians have been very reluctant to impose substantially lesser penalties on the perpetrators in the first – or third – sort of cases, as it would seem they logically should do. Perhaps the third case deserves reward.

There is plenty of evidence that by far the greatest harm to the greatest number of people is caused by politicians initiating wars of aggression with attacks on civilian targets, ruthlessly repressing dissidents within their own states and pursuing economic policies (nationally and internationally) that produce increased unemployment, inequality and poverty; by corporate executives killing their workers through inadequate attention to health and safety at work, polluting

the world with longlasting poisons and carcinogens, and marketing dangerous vehicles, drugs, foods, weapons and tools of torture and repression on a huge scale. Yet these people are seldom, if ever, prosecuted by the criminal justice system. Nor are libertarians typically very active or consistent in demanding any such prosecution.

On the other side of the coin, a small fine will cause much more harm to a poor person than to a rich one. Thus, the utilitarian Bentham argued for fines in proportion to the wealth of those fined, as well as to the extent of their wrongdoing.

We can extend this idea to consider also the prison system. There is reason to believe that some can survive and prosper in prison environments, largely at the expense of others, while those others can suffer appallingly. Psychopathic thugs outside remain psychopathic thugs inside, but inside there can be fewer obstacles to the exercise of their selfish violence. In the barbaric 'state of nature' prevailing in many prisons, the strongest and most unscrupulously violent (whether prisoners or warders) frequently rule, in the sense of effectively subjecting others to their will, in supplying money, drugs, labour and sexual services. The suffering of such others can, of course, be unimaginably terrible. But, again, libertarians have been reluctant to consider such real issues of the extent of pain and suffering of specific perpetrators, as, again, it would logically seem they should.

If we focus upon the utilitarian principle of pain just sufficient to ensure effective deterrence, other problems arise. The evidence indicates

> no consistent deterrent effect from punishment in society, and that the use of the penal system, and in particular imprisonment … is not an especially effective [individual] deterrent to crime. Recidivism rates remain high amongst those who receive custodial sentences [in one year, in England, approximately 50% of adults and 66% of juveniles were reconvicted within two years of a prior sentence]. And there is no evidence that longer custodial sentences produce better results than shorter sentences.[8]

Similarly, in Australia, 'on prison census night 2000, 56% of prisoners were known to have served a prison sentence on a previous occasion'.[9]

Of course, this does not tell us anything about all who might have avoided crime altogether through effective general deterrence by the prison system. But again, the evidence is that it is the 'subjective probability of being caught' as much as the nature of the punishment that is most significant in motivating those 'no longer fully anchored in the

value system of society, but who still have sufficient to lose from being caught and imprisoned'.[10] And detection rates for many crimes are widely known to be very low.

As Norrie says, 'to increase them would involve substantial increase in police powers and numbers and levels of state surveillance in society … with its own costs in terms of provoking other more serious kinds of crime',[11] including crimes of police corruption and public resistance. Certainly, libertarians and neo-liberals (generally) have problems in the increased tax costs and decreased individual freedom associated with such an expansion of state repressive powers.

Prisons

Libertarians have generally been strong supporters of maintaining and extending the prison system, without any major structural reform. Yet, as noted earlier, there are very serious problems with such a system.

Ever since the time of the reform movement (and before), prisons throughout the common law world have remained cultures of violence and disrespect, fostering and encouraging law-breaking rather than reducing it. As White and Perrone point out:

> An individual who enters the prison system undergoes a symbolic depersonalisation transition, stripped, probed, redressed and endowed with the status of a convict. As part of the process, the individual is required to take on the mores, customs and culture of the prison, all of which are premised upon conflict between inmates and guards.[12]

Even though the majority of prisoners have not been convicted of violent offences, prison officers perceive the prison situation as a battle for survival where they need to employ violent repression to control dangerous 'animals'. And 'the problems associated with an antagonistic prison environment and high levels of distrust and dishonesty make it very difficult for either side to change'.[13]

> The prison environment is far removed from the realities of outside life and this is a central factor in prisoners' difficulties of reintegration. The prison leaves indelible marks on the inmates, both in personal antagonism and frustration, and in the official blot on their record which will dog them the rest of their lives and severely effect their chances of successfully re-entering the maintenance of social life … The prison environment violates many of the known principles of social and psychological development. It promotes norms and practices which legitimate rather than reduce deviance.[14]

Cutting inmates off from friends and families, depriving them of heterosexual and caring relationships, depriving them of autonomy through subjection to a vast system of rigidly imposed rules, depriving them of all power over their own lives and circumstances – these are precisely the ways to undermine future coping and responsibility. Above all, throwing non-violent individuals together with others (including guards) with a history of violence in a situation effectively removed from all protection by the 'rule of law' is a way to create desperate, angry, bitter, physically and psychologically damaged people, rather than responsible, capable citizens. As Hogg points out, in Australia:

> the national rate of recorded prisoner on prisoner assaults is about two and a half times the rate of assault in the general community. If the ratio of unreported to reported assaults in the prison system were the same as that in the general community [about 4:1] the difference … would be about 11 times.[15]

At the same time, life within prisons 'is a judicial no-go area' with the courts deferring to executive authority in 'the management of the good order and discipline of the prison'.[16] Prisoners are subjected to continuous violence and unhealthy air and food, and forced to work for capitalist corporations with no proper payment, Workcover, occupational health or safety or workers' compensation coverage.[17] They have no protection from police authorities and effectively no legal rights to pursue any of these issues in the courts.

The criminal justice system in Australia systematically discriminates against particular social groups, with 50% of prisoners having failed to complete secondary schooling, 40% of females and 30% of males having long-term illnesses or disabilities, and one-third of males having been confined to juvenile institutions at some time. The unemployed are particularly likely to end up in gaol. And indigenous people are imprisoned at about 15 times the rate of non-indigenous people.[18]

With 25 per cent of young men in NSW gaols being raped, some on a daily basis, according to the research of NSW magistrate David Heilpern, with extensive use of hard drugs and rampant hepatitis C, with boys schooled in a life of crime and pleading with relatives to smuggle in drugs to stop them being bashed, there are very serious ethical questions of whether the existence of such institutions is not a much greater crime than the 'crimes' they are supposed to address, quite apart from their encouragement of further crimes in the wider

society and the general failure of deterrence. Here, indeed, we can add judges and prosecutors who send young cannabis users and fine defaulters to prison to be raped and murdered, to become addicted to heroin and infected with hepatitis C and AIDS, to the list of most serious harm causers.

The social-liberal response

At the end of the 19th century, with the birth of welfare state interventionism (in the context of international imperialist rivalries and related concerns about the unhealthy state of the working class undermining national competitiveness), we see the development of alternative social-liberal ideologies of sentencing, as a response to the failure of classical libertarian models to 'control crime or legitimate punishment'.[19] As Norrie says, such social liberals 'added to the ideological brew of utilitarian deterrence and retributive justice new antithetical views of the nature and role of punishment', specifically 'ideas of social control that took the form of 'welfare', a 'welfare sanction'[20] based upon ideas of treatment and training. As Norrie explains:

> central to these developments was the growth of the positivist school of criminology which developed in this period. The essence of this new 'science' was its claim to knowledge of the nature, causes and treatment of criminality and delinquency, which held out the possibility of arresting its development in the social body. Like … doctors [at that time arguing for the absence of free will in psychologically disturbed individuals] the criminologists rejected the classical ideas of the penal reformers concerning the responsibility and freedom of the individual.[21]

The new criminologists saw criminal behaviour as the product of particular social or physiological circumstances influencing individual psychology. And criminality became, for them, 'the target of direct technologies of behaviour control, rather than the respectful response to a pre-existing rational moral act'.[22] In particular, they saw imprisonment as an opportunity for 'stimulating or awakening the higher susceptibilities of prisoners, [developing] their moral instincts, [training] them in orderly and industrial habits',[23] rather than 'a sanction essentially rooted in deterrence' and pain infliction.

At the same time, for those few wrongdoers who proved impossible to reform owing to their irremediable damage, and who posed a continuing serious danger to the community (as identified by the appropriate experts), prisons also provided a means of community

protection through incapacitation. As White and Perrone point out:

> a quasi-military style of imprisonment adopted in the 1830s, with prisoners confined to small cells with strict rules of silence [supposedly maximising their opportunities to reflect upon past sins and thereby allowing their own consciences to contribute both to their punishment and spiritual healing], was replaced by the early 1900s with a more open sort of system, which allowed prisoners increased freedom of movement and interaction in exercise yards, and productive group work aiming to 'normalise' their behaviours and prepare their reintegration into society as productive and law-abiding citizens.[24]

However, it is again easy to see the problems with the new approach. Most obviously, there is the radical contradiction between continuing goals of retribution and deterrence (as previously understood), and new, humanitarian aims of moral reform and rehabilitative treatment through provision of an environment supportive of such 'positive' change in the offender. It is simply not possible (or, at least, extremely difficult) to seriously pursue the latter goals in the sort of environment generated and sustained by the former.

As White and Perrone note, the contradictions are further exacerbated when we factor in external pressures for 'cost effectiveness' as well, radically reducing the real material resources for effective rehabilitation, and contributing to overcrowding that itself further undermines such programmes.

As noted in the last section, despite rehabilitative ideas and interventions, prisons have generally remained cultures of violence, dishonesty and disrespect, which encourage crime, inside the prison itself and in the world outside. Such rehabilitative reforms as expanded education and skills training, drug treatment and group psychotherapy for prisoners have simply been grafted on to such a culture, without addressing the fundamental issues of conflict, brutality, dehumanisation and ill health considered earlier. And they have therefore never been given any real chance of general success.

Equally important, even an effective regime of genuine rehabilitation would fail to address major social-structural causes of crime. Effectively preparing offenders for constructive employment in the wider society counts for little if there are no jobs for them in that wider society. Treating them with care and respect while in confinement is of reduced significance if they are treated with a total lack of care and respect in the world outside, by employers, police and welfare services, for example.

Reaction

In this context we can understand the neo-liberal reaction against social-liberal, rehabilitative ideas in recent decades. On the one hand, this is merely part of the general shift away from social-liberal ideas arising out of the recession of the mid-70s and the subsequent rise to power of neo-liberal ideas, regimes and policies. But, at the same time, it has been possible for neo-liberals to point to the failure of previous regimes of rehabilitation in terms of continued high rates of recidivism and criminal activity.

So too have they focused upon criminological ideas of social causation of crime and social technologies of crime reduction, to argue that social-liberal reformers have treated criminals as mere objects of external manipulation, thus failing to respect their individual autonomy and powers of decision and choice. Genuine respect for such individuals requires treating them as free agents of their own actions, rather than mere 'puppets or 'effects' of external causes.

The truth of the matter is, as we have seen, that rehabilitation has seldom been given any real chance to succeed. In the few cases where genuinely rehabilitative policies have effectively replaced the culture of violence associated with retribution, as in the Dutch and Swedish prison systems, where there is no brutalisation of prisoners or stripping away of their rights and sentences are generally short, 'recidivism rates are no higher than in other countries with twice or three times the imprisonment rates'.[25]

There need be no necessary contradiction between ideas of social determinism and recognition of, and respect for, individual autonomy. As discussed in the last chapter, autonomy is a social creation: it depends upon social provision of material, psychological and emotional resources to enable individuals to make genuinely rational and responsible choices. Seen in this light, it is obvious that autonomy can be radically compromised by conditions that preclude or obstruct such rational and moral deliberation, choice and action.

Neither does rehabilitation need to treat offenders as mere objects of external manipulation and control. On the contrary, genuine 'rehabilitation' is a collaborative enterprise, integrally involving the understanding and active participation of the offender and aiming – precisely – to increase the scope of their effective 'choice' and rational decision-making (through offering, for example, new job opportunities).

Libertarian-inspired policies of mandatory sentencing and aboli-

tion of early-release schemes have increased prison overcrowding, leading to intensified conflict and violence within prisons and still further undermining limited rehabilitative efforts. Policies that have increased unemployment and/or inequality in the wider society have contributed to increased crime rates. More police and more prosecutions have further contributed to overcrowding, anger, frustration and violence.

This certainly looks like a general policy of neo-liberals in government to undermine social services, particularly through radical funding reductions with a view to creating chaos and scandal – which are built up by a sympathetic media. The regime then declares that such services are 'not working' and need to be abolished altogether.

In light of increasing corruption at all levels of public life in recent years, with politicians lying to the public on a daily basis and using their power to feather their own financial nests, and business people unconstrained by law or morality in their pursuit of ever greater personal wealth, rehabilitation as preparation for life outside of confinement does indeed become a problematic notion. Helping the generally poor and disadvantaged victims of the criminal justice system to function 'effectively' in this world would seem to require either helping them to prepare for reduced and miserable lives as docile drudge paupers in the service sector, or helping them to become more effective criminals on the other side of the fence – as 'managers'.

Socialist criminology

A socialist criminology focuses upon addressing what are seen as the real social causes of major suffering and privation. This means dismantling the structures of market relations and class power that (are seen to) produce greedy selfishness, inequality, poverty and discrimination and thereby contribute to those actions currently treated as crimes. It aims to achieve the abolition of criminal prosecution in a future communist society, through removing the social causes of crime.

During the process of social reconstruction, the approach of the transitional socialist state to selfish acts of social disruption would presumably be informed by social-liberal ideas of rehabilitative welfare, education and assistance, with therapeutic incapacitation of the small minority possibly unable to benefit from such assistance.

The typical libertarian and neo-liberal response to such ideas is to argue that it is human nature, rather than social relations, that is the

major cause of criminal wrongdoing. Crime, in the sense of selfish behaviour causing serious harm to others, will therefore continue to exist no matter what the form of prevailing social relations, and some sort of retributive punishment will still be required as both the morally respectful and the practically effective means of keeping criminal activity within acceptable bounds.

The argument of the chapter so far provides little support for any such – ultimately contradictory – libertarian ideas. Either social control (of harmful behaviour) is possible, in which case we must look to social insight, science and practical politics, rather than prejudice, to improve its effectiveness, or it is not, in which case there is even less excuse for gratuitous torture and cruelty. There is nothing remotely respectful about sending anyone into anything resembling the current (common law) prison system.

A wealth of everyday and historical experience tells us that treating people with (genuine) respect encourages them to treat others with respect. Giving them genuine non-criminal opportunities for exciting and rewarding and socially useful life activities, with fair and adequate recompense, is the best way to reduce criminal activity. And where they are unable to take advantage of such opportunities for one reason or another, comprehensive social welfare is crucial in protecting them from the need for criminal acts.

Human nature is a difficult and problematic concept. We have touched upon some of the difficulties in earlier chapters. At this stage, we merely note that human nature in the sense of hard-wired genetic propensities or predispositions or tendencies could, indeed, be a major cause of irresponsible and harmful behaviour in some sorts of social situation, where, for example, some innately programmed sequence of behaviour is triggered by some particular sort of environmental stimulus. But this in no way implies that such harm cannot be avoided by (a) removing such triggers or (b) rearranging circumstances such that the results of the behaviour in question are harmless, rather than harmful.

Discussion topics

1. What are said to be the purposes of punishment? Can these be justified?
2. Does the Australian prison system effectively fulfil these purposes? How might the situation be improved?

3 What should be the purposes of punishment? Could we do without it altogether?

Additional resources

D. Brown and M. Wilkie (eds), *Prisoners as Citizens: Human Rights in Australian Prisons*, Federation Press, Sydney, 2002.

R. White and S. Perrone, *Crime and Social Control: An Introduction*, Oxford University Press, Oxford, 1997.

Chapter 20

TERRORISM AND DEMOCRATIC RIGHTS

One of the most striking challenges for the 21st century is the protection of democratic rights. In the opening years of the new century, numbers of governments have used the threat of terrorism as a pretext to erode such vital principles as free speech, freedom of political association, prevention of arbitrary detention, and the right to seek asylum.

In Australia (and there are parallels elsewhere, notably in the United States and Britain), the dawn of the century saw three fundamental shifts in the state machinery with the introduction of legislation: in 2000 to permit the calling out of the military against civilian unrest; in 2001 to authorise the forcible turning away of refugee boats; and in 2002 to grant detention and proscription powers, as well as expanded surveillance powers, to the government and its security and intelligence agencies, notably the Australian Security Intelligence Organisation (ASIO).

These measures have profound implications for civil liberties, as well as for the future of international covenants, such as the Refugee Convention and the International Covenant on Civil and Political Rights. These global human rights instruments have proved largely irrelevant in curbing such powers (for the anti-refugee measures see chapter 21).

The Howard government in Australia followed the lead of the Bush administration in the United States and the Blair government in Britain by declaring that the September 11, 2001 terrorist attacks on the United States required an indefinite 'war' against terrorism abroad, accompanied by curtailment of legal rights at home.

Despite criticisms by civil liberties groups, both the British and American governments introduced severe anti-terrorism measures, including detention without trial and proscription of organisations.[1] Amnesty International condemned the Bush administration for breaching the International Covenant on Civil and Political Rights and other international protocols against arbitrary detention and inhuman treatment of prisoners.[2]

Significantly, the first two sets of Australian legislation pre-dated September 11, indicating that the anti-democratic trend is more fundamental than a response to the events in New York and Washington. Rather, these atrocities, and later the Bali bombing of 12 October 2002, were seized upon to retrospectively justify, as well as to introduce, new far-reaching alterations to the legal and constitutional framework.

These political and legal shifts, as this chapter will review, are profoundly anti-democratic. Despite the contrary impression created by the mass media, there is no evidence of strong popular demand for such measures; in fact, each legislative package aroused considerable public opposition. A careful review of the provisions and the circumstances in which they were introduced suggests that the purpose for their introduction is to strengthen the repressive capacities of the state against the free movement of people and other perceived threats to the political establishment.

Striking a balance?

It may be helpful to begin by examining the official justifications for the new measures. In an address to the Sydney Institute, delivered in the Mallesons Conference Room on 20 April 2004, Attorney-General Philip Ruddock outlined what he termed a 'new framework' for considering terrorism and the rule of law:

> The war on terror is like no other war in living memory. This is a war which may have no obvious conclusion, no armistice and no treaty. Victory in this war will not necessarily be measured by territory gained or regimes toppled. In this war victories will be measured by disasters averted and democracy strengthened. This war's victories will be measured by citizens feeling safe in their homes. This war's victories will be measured in the steadfastness and resolve of Australians to be cognisant of, but not to fear, a potential terrorist threat...
>
> Our Constitution, one of the world's oldest and most stable, provides us with a mechanism to protect our country and at the same time protect civil liberties through human security laws. In enacting such

> laws we are not only preserving traditional notions of civil liberties and the rule of law, but we are recognising that these operate in a different paradigm. If we are to preserve human rights then we must preserve the most fundamental right of all – the right to human security.

While insisting that the government is upholding the Constitution, civil liberties and the rule of law, the minister asserted that these now operate in a new paradigm: the right to human security, which is said to be the most fundamental right of all. Citing remarks by UN Secretary-General Kofi Annan, Mr Ruddock loosely defined 'human security' as encompassing human rights, good governance, access to education and health care and opportunities for individuals to fulfil their potential. All these, the minister asserted, depended upon a secure environment. Thus, in the name of defending civil liberties and the rule of law, they are said to no longer have any independent or absolute existence. Instead, they have been subsumed under another concept, human security. Making 'citizens feel safe in their homes' has become the chief criterion for the unknown duration of the current 'state of war'.

Despite the sweeping breadth of such claims, much of the debate and analysis in academic publications concerning the 'anti-terrorism' legislation begins with the proposition that a balance must be struck between 'national security' and 'civil liberties'. According to this approach, the only disagreement concerns where the balance should lie. By this measure, some inroads into civil liberties must be accepted. There is good cause to question this assumption, however.

In the first place, it can be argued that for powerful historical reasons, fundamental democratic rights should be regarded as absolute. They embody centuries of deep-going political struggles. In Anglo-Saxon law, civil liberties – such as habeas corpus, the presumption of innocence, the requirement of proof beyond reasonable doubt for a criminal conviction, freedom of association and free speech – were substantially forged in the conflict against the absolutist monarchy, from the Magna Carta of 1215 and culminating in the English Civil War of the 1640s and the so-called Glorious Revolution of 1688. Among these fundamental rights is freedom from detention without trial, as the US Supreme Court, by a 6–3 majority, commented in June 2004, in ruling that Guantanamo Bay detainees, including two Australians, David Hicks and Mamdouh Habib, could seek writs of habeas corpus in US courts. The majority judgment, delivered by Stevens J, suggested that at stake were democratic conceptions dating back nearly 800 years to the Magna Carta of 1215.

> Executive imprisonment has been considered oppressive and lawless since John, at Runnymede, pledged that no free man should be imprisoned, dispossessed, outlawed, or exiled save by the judgment of his peers or by the law of the land. The judges of England developed the writ of habeas corpus largely to preserve these immunities from executive restraint. [*Rasul v Bush*; *Al Odah v United States* (2004) 542 US (Cases No 03–343, 03–334) Quoting Jackson J in *Shaughnessy v United States ex rel Mezei*, 345 US 206, 218–19 (1953) (dissenting opinion)].

Second, there is much evidence to suggest that the 'war on terror' has been declared for definite political purposes, both foreign and domestic, rather than to protect the security of ordinary people. It is now widely acknowledged that all the claims made to justify the central international operation of the 'war' (the US-led invasion of Iraq), that is, the claims of 'weapons of mass destruction' and Saddam Hussein's links to terrorism, have collapsed. Moreover, whether or not the Bush administration knew in advance of plans for some kind of terrorist atrocity on September 11 – and that question still has to be answered – the outrages in New York and Washington provided the opportunity for the implementation of plans prepared much earlier – during the 1990s – for the conquest of Afghanistan and Iraq. The Middle East and Central Asia, as is well known, contain the largest proven concentrations of oil and natural gas reserves in the world. The US-led interventions in the region and the establishment of US military bases throughout Central Asia have added weight to the evidence that Washington's underlying ambition is to secure hegemony over this entire vital expanse.

There is equal reason to doubt the domestic side of the war. Since the September 11, 2001 attacks, the Howard government has eroded longstanding legal and democratic rights in the name of combating terrorism. Yet, on the face of it, none of ASIO's new powers were necessary to protect ordinary people against terrorism. Many submissions to parliamentary committees inquiring into the proposed legislation, including those of the Law Council of Australia and the Civil Liberties Councils of New South Wales and Victoria, questioned the need for the entire package. As pointed out by a parliamentary library report, any conceivable terrorist activity, such as murder, bombing, hijacking, kidnapping and arson, was already a serious crime under existing law.

ASIO hardly needed new powers to detect terrorists. It already had a vast array of powers to tap phones, install listening devices in offices and homes, intercept telecommunications, open people's mail,

monitor online discussion, break into computer files and databases, seize computers and use personal tracking devices. The ASIO Director-General or his delegated officers could already issue emergency search and entry warrants, allowing officers wide scope to conduct operations against political activists and organisations, as well as to infiltrate them. Moreover, ASIO is part of an extensive security and intelligence network, which incorporates the external Australian Secret Intelligence Service, the Office of National Assessments, special state police units (formerly called Special Branches), the military's Joint Intelligence Office, the Defence Intelligence Organisation, the Defence Imagery and Geospatial Organisation and an electronic eavesdropping agency, the Defence Signals Directorate.

The military call-out legislation

In September 2000, amid considerable public controversy, the Australian Labor Party (ALP) joined forces with the Government of Prime Minister John Howard to have military call-out legislation passed through both houses of the Commonwealth parliament.[3]

Both the Government and the Opposition declared that it was necessary to have the legislation in place before the Sydney Olympic Games. In the brief parliamentary debate, references were made to the need to counter possible terrorism at the Olympics, where some 4 000 military personnel were deployed.[4] After expedited examinations by two Senate committees, whose recommendations for minor amendments were partially adopted,[5] the legislation was ultimately passed on the last day of sitting before the opening of the Games. Despite this haste, the Act was not invoked during the Olympics.[6] This suggests that the Olympics merely provided a pretext for the legislation, which in fact has more underlying purposes.

Under the amended *Defence Act 1903*, the federal government has the power to call out the armed forces on domestic soil against perceived threats to 'Commonwealth interests', with or without the agreement of a state government. Once deployed, military officers can order troops to open fire on civilians, as long as they determine that it is reasonably necessary to prevent death or serious injury. Soldiers will have greater powers than the police in some circumstances, including the right to shoot to kill someone escaping detention, search premises without warrants, detain people without formally arresting them, seal off areas and issue general orders to civilians.[7]

The legislation authorises the prime minister, the Defence minister and the attorney-general to advise the governor-general (the commander-in-chief of the armed forces under the Australian Constitution) to call out military personnel to deal with 'domestic violence'. The term 'domestic violence' does not correspond to the modern sense of the phrase, which refers to violence within homes or families. It is a vague expression, undefined legislatively or judicially, found in s 119 of the Australian Constitution, which provides that 'the Commonwealth shall protect every State against invasion and, on the application of the Executive Government of the State protect such State against domestic violence'. Unlike s 119, however, the new provisions do not require any invitation from a state government before troops are called out.

Both the Government and the Labor Party proposed limited amendments in an effort to meet certain objections from some state governments and to head off public concern about the impact on civil liberties, but the legislation's essential content remained the same: to authorise the use of the military to deal with civilian disturbances, including political and industrial unrest. The fact that such legislation was passed suggests a bipartisan expectation in official political circles that, in the coming period, troops will be required to deal with disturbances that the police forces cannot contain.

Historical context

For more than 100 years, the domestic deployment of troops has been politically contentious and clouded by legal uncertainties. In the words of one author, although Australia was established as a penal colony under military administration, 'with the passage of time, the evolution of the Australian political system ensured a clear distinction between military powers and civil powers'.[8] During the 19th century, martial law was declared several times to deal with riots and rebellions, but the last clear exception to the military–civil division of power occurred in 1891 when the Queensland government used troops to help the police suppress a sheep shearers' strike.[9]

This division of power was enshrined in the Australian Constitution at Federation in 1901. The military power was handed to the Commonwealth under s 51(xxxi), the colonial defence forces were transferred to the Commonwealth by s 69, and under s 114 the states were forbidden to raise military or naval forces without the

consent of the Commonwealth parliament. Residual authority over domestic law and order remained in the hands of the states and their police forces.

The constitutional demarcation became embedded in public consciousness. Domestic use of the armed forces was widely regarded as conduct to be expected of a military or autocratic regime, not a democratic government. On the only occasion since Federation that a Commonwealth government called out the military in an urban situation – following a bomb blast outside a regional Commonwealth heads of government meeting at the Sydney Hilton Hotel in 1978 – the sight of armed soldiers patrolling highways and the streets of the NSW town of Bowral caused public consternation.[10]

The legislation challenges a political and legal tradition opposing the use of the military to suppress domestic unrest – a principle that dates back to the 17th-century struggles against the absolutist monarchy in Britain. In the lead-up to the English revolution of the 1640s, the 1628 Petition of Right demanded that Charles I remove the 'great companies of soldiers and mariners [who] have been dispersed into diverse counties of the realm … against the laws and customs of this realm and to the great grievance and vexation of the people'. The petition is regarded as making it unconstitutional for the Crown to impose martial law on civilians.[11] As a result of the 1688 settlement between the monarchy and the parliament, the Bill of Rights declared it illegal for the Crown to raise or keep an army without parliamentary consent.[12]

There has been no recorded case of martial law in Australia since Federation in 1901 but it was invoked several times during the 19th century to suppress convicts, Aborigines and workers. The strike struggles of the 1890s saw troops mobilised against specific demonstrations and gatherings, with orders to shoot to kill strikers and their supporters. In one infamous incident, Colonel Tom Price issued the following instruction to a volunteer unit during the extended Australian maritime strike of 1890:

> [I]f the order is given to fire, don't let me see any rifle pointed in the air; fire low and lay them out so that the duty will not have to be performed again.[13]

The turmoil of the 1890s led to s 119 being inserted in the Constitution, to allow the military to be mobilised against an 'uncontrollable situation'.[14] The expression 'domestic violence' was bor-

rowed from Article IV of the US Constitution, s 4 of which specifies that the United States shall protect each state, on the application of its legislature, against 'domestic violence'. The statutory embodiment of this provision in 10 USC § 331 (1964) uses the more specific term 'insurrection', suggesting that an extremely serious level of rebellion must be involved – one that threatens the very existence of a state government.

In the early years of the 20th century, Australian state governments requested military intervention on at least six occasions, to deal with such anticipated incidents as 'general strike riot and bloodshed', 'disturbances', wharf strike 'violence', 'labour troubles' and the 1923 Victorian police strike. On each occasion, it seems, the federal government declined on the basis that the state police were capable of dealing with the threat (although troops were sent to guard federal buildings, including post offices, during the Victorian police strike). Only one of those requests – by Queensland in 1912 – was formally made under s 119. Thus, s 119 has never been applied.

The legislation

Once deployed under the amended *Defence Act*, the military forces will have wide-ranging powers that they currently do not have in civilian situations. The most revealing measures are those contained in s 51T on the use of 'reasonable and necessary force'. Soldiers will be permitted to cause death or grievous bodily harm where they believe 'on reasonable grounds' that such action is necessary to protect the life of, or prevent serious injury to, another person – including the soldiers involved. A person 'attempting to escape being detained by fleeing' may be killed or caused grievous bodily harm if they have been called on to surrender and a soldier believes on reasonable grounds that the person cannot be apprehended in any other way.

Confronted by public hostility to its earlier unconditional endorsement of the Act, the ALP moved an amendment forbidding troops to 'stop or restrict any protest, dissent, assembly or industrial action, except where there is a reasonable likelihood of the death of, or serious injury to, persons'. The Government added a final clause, 'or serious damage to property', which Labor accepted.

The resulting s 51G opens the way for wide use of the call-out power. Likelihood of property damage can be alleged easily. As Independent MP Peter Andren put it, 'a rock thrown through the front

door of the Crown Casino [the venue of the 2000 World Economic Forum in Melbourne] could give rise to such a call-out'.[15] As for the likelihood of injury, that could be created by a police attack on demonstrators.

Doubts remain about the constitutional validity of the provisions, notwithstanding the fact that a number of authorities have taken a generous view of the Commonwealth's powers to call out the military.[16]

It is clear that the implications of the legislation go far beyond the Sydney Olympics. The Government and the ALP Opposition rejected amendments to insert a sunset clause that would revoke the legislation after the games. One can only conclude that the Act has given effect to a permanent shift in the military's role. Australia's constitutional and legal framework has been altered to allow for military intervention to deal with any potentially destabilising internal unrest or political dissent.

'Counter-terrorism' laws

In June 2002, the Howard government, again assisted by the ALP, secured passage of a raft of 'counter-terrorism' bills, handing unprecedented powers to the executive government and the intelligence and police agencies. The bills introduce sweeping definitions of terrorism and treason, both now punishable by life imprisonment, which could outlaw many forms of political protest and industrial action. They contain powers to ban political parties, freeze their funds and jail their members for alleged support of 'terrorism'. In addition, they reverse the burden of proof for a range of serious offences, effectively requiring defendants to prove their innocence.[17]

Because of public opposition and adverse parliamentary committee reports, the government temporarily withdrew one measure, the Australian Security Intelligence Organisation Legislation Amendment (Terrorism) Bill (ASIO Detention Bill), which sought to authorise ASIO to detain people without charge and interrogate them incommunicado. With Labor's support, the Bill was finally passed in June 2003 with limited amendments, along the lines proposed by parliamentary committees.

Context and pretext

It is ironic that 50 years after the *Australian Communist Party Case* (*Australian Communist Party v Commonwealth* (1951) 83 CLR 1) and the subsequent defeat of a referendum to ban the Communist

Party, the main political parties passed legislation that goes beyond the measures of 1950–51 in its potential to outlaw dissenting political activity. The Cold War, which provided the setting for the Menzies government's proposals, has ended, but, instead of a new period of political freedom, we are witnessing far-reaching moves against traditional democratic norms.

The Howard government rejected previous advice, adhered to by successive administrations since the Hilton Hotel bomb blast, that it was unnecessary, inadvisable and constitutionally questionable to introduce generic anti-terrorism laws. In the 1979 *Protective Security Review Report*, Justice Robert Hope, while recommending a major boost to the powers and resources of the police, intelligence and security forces, did not recommend the creation of new criminal offences, stating: 'Terrorism by its nature involves breaches of the ordinary criminal law.' In an opinion commissioned by the Fraser government as part of Justice Hope's review, former High Court Justice Victor Windeyer came to the same view.

One must ask why the government has now overturned these precepts. The legislation punishes violent or other criminal activity far more severely if offenders are motivated by political, religious or ideological considerations than if they are acting for revenge, rage, greed, lust or other motives. This indicates that it is political motives, rather than the conduct itself, that the government is seeking to punish. This suggests that the 'war on terrorism', like the 'war on communism' a half-century ago, is being used for political ends.

Certainly, the Howard government's rhetoric is reminiscent of the campaign waged a half-century ago. After winning the 1949 election in the wake of the coal miners' strike, Prime Minister Robert Menzies claimed a 'political mandate' to place Australia on a 'semi-war footing' against communism. Against a backdrop of global anti-communism, the Communist Party Dissolution Bill was the incoming government's first piece of legislation. The Bill's recitals claimed that its measures were required for the 'security and defence of Australia' in the face of a dire threat of violence, insurrection, treason, subversion, espionage and sabotage.

The High Court of Australia, however, rejected the use of these recitals to validate the government's claim to be exercising the defence, incidental and executive powers of the Commonwealth. The judges warned of the corrosive dangers of unfettered executive power. Dixon J stated:

> History, and not only ancient history, shows that in countries where democratic institutions have been unconstitutionally superseded, it has been done not seldom by those holding the executive power. Forms of government may need protection from dangers likely to arise from within the institutions to be protected.[18]

The court's stance was, in effect, vindicated by the defeat of the 1951 referendum. It remains to be seen whether today's High Court will restate these principles, if and when the current legislation is challenged.

Terrorism, treason and espionage

Central to the legislative package are far-reaching definitions of terrorism, treason and espionage. These offences have become some of the most serious on the statute books, with severe penalties. The first two are punishable by life imprisonment; the third by 25 years' imprisonment. Under the Security Legislation Amendment (Terrorism) Bill, terrorism extends to acts or threats that advance 'a political, religious or ideological cause' for the purpose of 'coercing or influencing by intimidation' any government or section of the public. 'Advocacy, protest, dissent or industrial action' is exempted but not if it involves harm to a person, 'serious damage' to property, 'serious risk' to public health or safety, or 'serious interference' with an information, telecommunications, financial, essential services or transport system.

The legislation imposes jail terms ranging from life to 10 years for preparing, planning or training for 'terrorist acts' and for possessing documents or other objects used in the preparation of such acts. A person can be jailed for possessing such a 'thing' even if they did not know it was used for terrorist purposes, but were merely 'reckless' as to that fact.

This definition could cover any demonstration or strike action in which a person was injured or felt endangered, given that the purpose of many protests and strikes is to apply pressure to a government, employer or other authority. Nurses taking strike action that shuts down hospital wards in support of a political demand for greater health spending, for example, could be accused of endangering public health and thus be charged as terrorists.

The various, related, terrorist offences could apply to a wide range of political activity, such as planning or participating in a protest outside government buildings or facilities where damage is alleged to have occurred. Demonstrators who block roads or entrances to financial institutions, such as the stock exchange, could be charged as terrorists, as could computer hackers.

During questioning in a Senate committee hearing, the attorney-general's representatives admitted that someone who cut through a fence at the Easter 2002 protest at the Woomera refugee detention centre or who invaded the parliament building during a 1996 trade union rally could have been charged with terrorism. The officials acknowledged that a picketing striker who caused property damage or a person who possessed a mobile phone used to discuss a violent act could be prosecuted under the new provisions.

While citing the September 11 attacks in the United States as its justification, the Howard government has adopted a definition of terrorism that goes beyond the Bush administration's USA Patriot Act, which covers activity that is dangerous to human life and violates existing criminal laws. The Howard government's version is based on the British Blair government's *Terrorism Act 2000*, but goes further in specifying disruption to various communications systems.

Power to outlaw organisations

Under the original version of the Terrorism Bill, the Attorney-General could have proscribed any organisation on a number of vague grounds, including his view that a group had 'endangered, or is likely to endanger, the security or integrity' of Australia or another country. In the light of the wide meanings that can be given to the term 'national security' and the difficulties of obtaining judicial review of its use by government and intelligence agencies, these criteria opened up wide scope for political abuse.

In the face of public opposition, the government was compelled to back down on its original proposal. But the amended version allows the government to issue regulations to outlaw parties or groups if the UN Security Council has listed them as terrorist. A court can also declare an organisation to be 'terrorist'.

Proscription orders may have far-reaching implications. Any person who directs or provides support to the activities of a terrorist organisation, knowing it to be terrorist, can be jailed for 25 years or, if they are 'reckless' as to whether the organisation is terrorist or not, for 15 years. A member of a group banned under a regulation faces up to 10 years' imprisonment. Membership is defined to include 'informal membership' or taking 'steps to become a member'. It is a defence to have taken 'reasonable steps' to cease membership 'as soon as practicable' after knowing the organisation was terrorist, but the burden of proof lies on the defendant.

The legislation also retains a backdoor method for banning organisations by freezing their funds, even if they have not been formally declared terrorist. The attorney-general can freeze assets or proscribe groups if a UN Security Council freezing order has been issued. Either the minister can 'list' an organisation by Gazette notice or the governor-general may make proscription regulations. Anyone collecting or providing donations for the organisation can be jailed for five years. If the funds are used for terrorist purposes, the penalty is life.

Detention without charge

Following the ultimate passage of the ASIO Detention Bill, ASIO has the power to detain and question people without charge or trial. ASIO and Federal Police officers can raid anyone's home or office, at any hour of the day or night, and forcibly take them away, interrogate and strip-search them and hold them incommunicado, effectively indefinitely through the issuing of repeated warrants.[19]

Detainees did not need to be suspected of a terrorist offence, or any other criminal offence. The attorney-general can certify that their interrogation would 'substantially assist the collection of intelligence that is important in relation to a terrorism offence', even if no act of terrorism has occurred. This power could be used to detain journalists and political activists, as well as the children, relatives or acquaintances of supposed terrorism suspects. Any detainee who refused to answer ASIO's questions would be liable to five years' imprisonment.

Those detained have no right to know why they are being hauled off for interrogation. If they resist, violent force, including lethal force, can be used against them. If they refuse to answer any question or hand over any material that ASIO alleges they possess, they face five years' jail. Detainees, including teenagers as young as 16, will be unable to contact their families, friends, political associates or the media. If they know the name of a lawyer, they can contact them for legal advice, but only if ASIO does not object to the lawyer.

Even if ASIO accepts a detainee's choice of lawyer, questioning can commence without the lawyer being present. In any case, the lawyer cannot object or intervene during questioning – if they do, they can be ejected for 'disrupting' ASIO. If they inform a detainee's family or the media about the detention, they too face up to five years in jail. A lawyer who is provided information by a client may also be detained for interrogation. The Act does not protect legal professional privilege in communications between lawyer and client.

Initial detention can last for up to seven days, including three eight-hour blocks of questioning over three days, but the attorney-general can easily approve further seven-day periods. To justify serial extensions, ASIO and the government simply have to claim that 'additional to or materially different' information has come to light.

In a significant departure from established law, the Act effectively reverses the burden of proof, overturning a basic protection against police frame-up. If ASIO alleges a person has information or material, the onus is on the individual to prove otherwise.

Section 34JB permits police officers to use 'such force as is necessary and reasonable' in breaking into premises and taking people into custody. This clause gives police the power to kill or cause 'grievous bodily harm', as long as they believe it necessary to protect themselves. In addition, officers may use 'reasonable and necessary' force to conduct strip-searches.

Interrogation must be videotaped and conducted in the presence of a 'prescribed authority', that is, a judge, retired judge or presidential member of the Administrative Appeals Tribunal. Studies have shown that videotaping of questioning, currently required for police questioning in most Australian jurisdictions, is no guarantee against the planting of evidence and extraction of false confessions. And a government can readily appoint retired judges or tribunal members, with no judicial tenure, who may be amenable to its requirements.

The legislation has radically extended ASIO's powers. The agency previously had no powers of arrest or interrogation. The state and federal police can detain people, but only on suspicion of committing a criminal offence, and those suspects must be either charged or released within a short period and generally cannot be detained for interrogation. Prisoners have the right to legal counsel, who can be present during questioning, and to remain silent. With the notable exception of the detention of asylum seekers, detention without trial is regarded as unconstitutional. A citizen is entitled to decline a request to attend a police station 'to assist police'.

The new powers are unparalleled. Not even during two world wars did an Australian government seek to overturn freedom from arbitrary detention (with the controversial exception of rounding up people of German, Japanese and Italian origin as 'enemy aliens') or to abolish the centuries-old common law right to remain silent.

Despite the level of public disquiet, in April 2002 the leaders of the Australian states and territories, all run by Labor governments, agreed

at the Conference of Australian Governments summit to formally refer their powers over terrorism to the federal government. Their decision has the potential to give the Commonwealth substantially unfettered law-making and police enforcement power over politically related crime for the first time since Federation in 1901, possibly freeing the Howard government of the need to find precise constitutional heads of legislative or executive power for its measures.

Even so, the potential constitutional problems with the laws are manifold, arising from the Commonwealth parliament's lack of general power to legislate with respect to criminal law or 'terrorism', as well as the implied right under the Constitution to political communication and, arguably, freedom of association. In addition, detention without trial may infringe on judicial power and the separation of powers.

In order to secure passage of the ASIO Detention Bill the government agreed to insert a three-year sunset clause. However, the government made plain its intention to seek the legislation's renewal and has introduced amendments to strengthen ASIO's detention powers.

ASIO operations cloaked in secrecy

Amendments passed in December 2003, again supported by the Labor Party, effectively gag all public protest against, or even reporting of, the use of the new powers. It is now a crime, punishable by up to five year's jail, to publicly mention any operation involving ASIO's powers. Even if ASIO itself breaks the law, for example, by detaining someone for more than seven days without obtaining a new warrant, any journalist who reports the case could be imprisoned.[20]

In effect, these measures outlaw political campaigns against arbitrary or illegal detentions. If someone sees a person being hauled away by ASIO or federal police for questioning, they cannot disclose that fact to anyone – not even a family member, friend, civil liberties group, member of parliament or political party. If a detainee's family or associates somehow find out about the detention, they cannot publicly comment on it in any way.

The ASIO detention laws passed earlier already prohibited detainees or their lawyers from alerting their families, the media or anyone else that they had been detained. This gag has been broadened to cover all people, not just detainees and lawyers, and extended for the full 28-day period of a warrant.

A further two-year prohibition was imposed on the public disclosure by anyone of 'operational information' that was obtained, directly or indirectly, from the questioning process. 'Operational information' is defined in the widest possible terms. It covers all ASIO information, sources of ASIO information and any 'operational capability, method or plan' of ASIO.

The government insisted on pushing these provisions through parliament unamended, despite strong protests from civil liberties and media organisations. Amnesty International declared: 'The level of secrecy and lack of public scrutiny provided for by this Bill has the potential to allow human rights violations to go unnoticed and in a climate of impunity.'[21] Liberty Victoria stated: 'These secrecy offences pose a grave threat to Australia's democracy and could enable the government of the day to impose a "war of terror" against its political opponents or vulnerable sections of the community.'[22]

Australia's main media proprietors' groups – Fairfax, News Ltd, SBS, the ABC, the Australian Press Council and Commercial Radio Australia – warned: 'This has the potential to completely remove from public scrutiny all discussion of ASIO's activities in relation to terrorism.'[23]

This legislation represents a grave erosion of political and press freedom. Even if ASIO itself breaks the law, for example, by detaining someone for more than seven days without obtaining a new warrant, any journalist who reports the case could be imprisoned. In effect, these measures outlaw political campaigns against arbitrary or illegal detentions. If someone sees a person being hauled away by ASIO or federal police for questioning, they cannot disclose that fact to anyone – not even a family member, friend, civil liberties group, member of parliament or political party. If a detainee's family or associates somehow find out about the detention, they cannot publicly comment on it in any way.

It is now possible for ASIO to cloak virtually all its operations in secrecy, simply by obtaining a questioning warrant from the attorney-general. For that reason alone, the latest legislation increases the danger that ASIO's detention powers will be abused. ASIO has a long record in this regard. Since the Chifley Labor government established the intelligence service in 1949, it has been used by successive governments, Labor and conservative alike, to monitor, disrupt and harass a wide range of political opponents, including Labor Party members, trade unionists, anti-war activists, students and socialists.[24]

Conclusion

Taken together, these laws represent a grave threat to basic democratic rights. Serious inroads have been made into longstanding principles such as no detention without trial, the presumption of innocence and freedom of speech and association. The pressures of globalisation and the 'war on terror' have set the stage for measures that substantially expand the powers of the military and security agencies. Both the context of the legislation and the extraordinary reaches of its measures invite constitutional challenge, as well as public opposition.

Discussion topics

1. What is terrorism? How has it been defined officially?
2. How does it differ from state terrorism?
3. It is possible, or appropriate, to strike a 'balance' between the 'war on terrorism' and basic democratic rights?

Additional resources

'Amnesty International's Concerns Regarding Post September 11 Detentions in the USA', Amnesty International, March 2002.

A. Bacevich, *American Empire: The Realities and Consequences of US Diplomacy*, Harvard University Press, Cambridge, 2002.

Call Out the Troops: An Examination of the Legal Basis for Australian Defence Force Involvement in 'Non-defence' Matters, Parliamentary Library, Research Paper No 8, 1997–98, Parliament of Australia, Canberra.

B. Graham and J. Nussbaum, *Intelligence Matters: The CIA, the FBI, Saudi Arabia, and the Failure of America's War on Terror*, Random House, New York, 2004.

N. Hancock, *Terrorism and the Law in Australia: Legislation, Commentary and Constraints*, Parliamentary Library, Research Paper No 12, 2001–02, Parliament of Australia, Canberra.

N. Hancock, *Terrorism and the Law in Australia: Supporting Materials*, Parliamentary Library, Research Paper No 13, 2001–02, Parliament of Australia, Canberra.

M. Head, 'Another Threat to Democratic Rights: ASIO Detentions Cloaked in Secrecy' (2004) 29 *Alternative Law Journal* 127.

M. Head, '"Counter-terrorism" Laws: A Threat to Political Freedom,

Civil Liberties and Constitutional Rights' (2002) 26 *Melbourne University Law Review* 266.

M. Head, 'The Military Call-Out Legislation – Some Legal and Constitutional Questions' (2001) 29 *Federal Law Review* 1.

M. Head, 'Olympic Security: Police and Military Plans for the Sydney Olympics – A Cause for Concern' (2000) 25 *Alternative Law Journal* 131.

J. Hocking, *Beyond Terrorism: The Development of the Australian Security State*, Sydney, Allen & Unwin, 1993.

H. Lee, *Emergency Powers*, Lawbook Co, Sydney, 1984.

H. Lee, P. Hanks and V. Morabito, *In the Name of National Security: The Legal Dimensions*, Sydney, LBC, 1995.

Royal Commission on Australia's Security and Intelligence Agencies (Commissioner Justice Robert Hope), *Report on the Australian Security Intelligence Organisation*, AGPS, Canberra, 1985.

Royal Commission on Intelligence and Security (Commissioner Justice Robert Hope), *Fourth Report: Volume 1*, Commonwealth Government Printer, Canberra, 1978.

R. Ruddock, 'A New Framework: Counter Terrorism and the Rule of Law', address to Sydney Institute, 20 April 2004, accessed 29 June 2004, <http://152.91.12/www/MinisterRuddockHome.nsf/Alldocs/RWPB046617DB0869>.

V. Windeyer, 'Certain Questions Concerning the Position of Members of the Defence Force When Called Out to Aid the Civil Power', in R. Hope, *Protective Security Review Report*, AGPS, Canberra, 1979, Appendix 9.

Chapter 21

REFUGEES AND THE NATION-STATE

It has been apparent for some years that the global refugee system is in grave crisis. According to the available statistics, the flight of people from their countries of birth grew dramatically in the final two decades of the 20th century and this mass movement is likely to grow in the 21st century. Increasingly, they are resorting to unauthorised methods of entry, often at great risk to their lives.

Many western governments are having considerable difficulties, logistically, diplomatically and politically, in removing those denied refugee status. Governments are spending mounting sums on detecting and detaining unwanted arrivals, deciding their fate and administering the outcomes, while giving decreasing funds to the UN High Commissioner for Refugees (UNHCR), which is responsible for most of the world's displaced persons. In 2000, a UNHCR-commissioned analysis of European governments' responses to the growth of 'people smuggling' concluded that official policy risks 'ending the right of asylum in Europe' and that the 'current status quo is practically and ethically bankrupt from all positions'.[1]

In Australia, governments have taken the policy of seeking to block and deter unwanted arrivals to the extent of detaining asylum seekers, usually in remote semi-desert locations, and, since the Tampa affair of August 2001, militarily barring entry to refugees.[2] Severe police and security methods, including the use of mass arrests, water cannon, tear gas and solitary confinement, have failed to quell the unrest in the detention camps – expressed in hunger strikes, mass breakouts and determined protests.

This chapter reviews the development of the refugee debate, examines the basic contradictions and fundamental flaws of the 1951 Convention Relating to the Status of Refugees (Convention), considers the current debate over the future of the Convention's future, and argues for a new paradigm that challenges the dichotomy between refugees and immigrants and recognises the essential democratic right to travel and live where one chooses.

How the refugee numbers have grown[3]

In 1951, when the UNHCR was established, there were an estimated 1.5 million refugees worldwide. On 1 January 2000, the UNHCR considered 22.3 million people to be 'of concern'. They included 11.7 million refugees, 1.2 million asylum seekers, 2.5 million repatriated refugees and 6.9 million internally displaced persons and others of concern. Another 13–18 million internally displaced persons were outside the UNHCR's jurisdiction, as were an estimated 3.5 million Palestinians. This gives a total of 43 million.

The number of people 'of concern' to the UNHCR nearly doubled during the 1990s, from 14.9 million in 1990, reaching an all-time high of 27 million in 1995, in the wake of the first Gulf War against Iraq, the fomenting of communalism in Yugoslavia, and the eruption of ethnic warfare in Rwanda and Central Africa's Great Lakes region.

Judging by the UNHCR's statistics for 'asylum applications lodged in selected countries', the rise in the 1990s followed an even greater increase during the 1980s. That figure jumped from 180 000 in 1980 to 435 000 in 1989, before skyrocketing to almost 850 000 in 1992 and ending the decade at 538 000 in 1999.

Of the largest concentrations of refugees in 1999, 3.1 million came from Afghanistan and Iraq, about 2.7 million from sub-Saharan Africa and nearly 800 000 from Bosnia and Croatia. Most are living in camps in neighbouring countries, usually among the poorest in the world.

The rise in refugee numbers has been related to definite economic and political processes, particularly the collapse of the Stalinist-ruled states in Eastern Europe from the late 1980s, the consequent end of the Cold War, the US-led bombing of Iraq and Serbia and the outbreak of regional conflicts in Eastern Europe, the Middle East, Asia and Africa. In Europe, thousands took the opportunity to flee from the Stalinist regimes and tens of thousands more sought refuge from the

severe social dislocation and civil and ethnic conflict associated with the restoration of capitalism. With the imposition of International Monetary Fund restructuring measures in one country after another, this global movement is likely to grow.

More fundamental driving forces are also at work. The increased demand for asylum has occurred amid an unprecedented globalisation of the world economy since the mid-1980s, creating massive flows of international capital, the rapid shift of production processes from country to country, and a worldwide labour market. At the same time, the ever widening gulf between the capital-rich, technologically advanced and militarily powerful countries and the rest of the world has fuelled the demand for the right to escape poverty.

Moves to restrict the right to asylum

While embracing the global restructuring of economic life, western governments have generally sought to erect new barriers to the movement of ordinary working people. One British writer has suggested that 'for a growing list of governments the best interpretation of the Convention Relating to the Status of Refugees can only be to run it through the shredder'.[4]

Some of the arguments raised by those calling for the Convention's restriction cannot withstand careful scrutiny. There is official opposition to the growth of 'human trafficking'. By this standard, those who sought to assist Jews escape the Nazis would have been damned as people smugglers. There is no doubt that so-called people smuggling has become a big business. By one estimate, one million people were transported in illegal operations worth up to $US20 billion in 1999.

But if refugees are seeking assistance from private sources, and often facing life-threatening conditions as a consequence, this is largely due to the steady erosion of refugee protection over the past decade. As governments have restricted entry and intensified measures to detect and exclude unwanted arrivals, refugees have made ever riskier efforts to gain a safe haven.

Similarly, the official condemnation of 'queue jumping' is questionable. Far from being in an orderly waiting list, those seeking safe haven confront impossible situations and terrible delays. Those likely to be the most needy – refugees in Africa, Asia and the Middle East – are the least likely to be accepted. To take the case of Australia, out

of the 7500 places for offshore applicants in 2000, 45 per cent were given to Europeans, leaving 2206 places for the entire Middle East and 1738 for all of Africa. On average, applications to Australia take 18 months to come up for consideration.

Moreover, the Australian government cuts its 12 000 per year quota of humanitarian and protection visas for offshore applicants by the number of asylum seekers who reach Australia independently and are granted refugee status. This policy pits the two groups – both in urgent need of protection – against each other.

The charge that asylum seekers who arrive without permission are 'illegal' entrants is equally dubious. In most cases they have broken no laws, and have not been convicted of any offence. In Australia, they are detained without trial. In any case, because refugees, by necessity, are often forced to escape from their countries and mislead authorities, Article 31 of the Convention stipulates that governments should not penalise applicants 'on account of illegal entry or presence'.

Another common argument is that a distortion of priorities is created when western governments spend far more on processing asylum applications than they donate to the world refugee effort. In 2000, Britain spent more on dealing with asylum seekers – $US2.2 billion – than the UNHCR budget of $1.7 billion. Australia spends as much each year on the Refugee Review Tribunal, just one level of the determination process, as it allocates to the UNHCR. Yet, this disparity underscores the paltry sums that western governments give the UNHCR. Donations to the UNHCR have decreased and its total income of some $US700 million in 2000 fell far short of the budgeted $1.1 billion.

The distortion of priorities is primarily a product of the measures being taken to undermine the right to asylum. One of the UNHCR's leading officials has summed up the situation as follows:

> Broadly speaking, two parallel trends have emerged, both of which have impacted negatively on the accessibility of asylum and the quality of treatment received by refugees and asylum seekers. The first has been the growth in an overly restrictive application of the 1951 Refugee Convention and its 1967 Protocol, coupled with a formidable range of obstacles erected by states to prevent legal and physical access to their territory. The second is the bewildering proliferation of alternative protection regimes of more limited duration and guaranteeing lesser rights than those contained in the 1951 Convention.[5]

Basic contradictions

Governments have generally facilitated the globalisation of economic life, seeking to attract wealthy international investors, while shutting their doors to the impoverished. They have deregulated their financial markets to enhance the movement of capital, but denied the same freedom to labour. These responses point to several basic contradictions. These can be summarised as follows:

- nation-state versus global economy;
- mobility of capital versus border controls on people;
- one law for the wealthy and another for the poor.

Nation-state versus global economy

The growing global mobility of people and commerce is increasingly in conflict with the efforts of governments to prevent the flow of unwanted arrivals. There is a worldwide movement of people, whether for employment, business, education, tourism, sport, entertainment, scientific and cultural exchange or social and political intercourse. More than ever, we live in a global village, linked electronically and by air travel. In large measure, this is the inexorable product of globalisation. National borders are becoming increasingly anachronistic.

The world economy today is characterised by the daily movement of vast quantities of capital across national borders, as international financial institutions scour stock and bond markets for the highest return on their investments. In the decade 1980–90, the volume of cross-border transactions in equities grew at a compound rate of 28 per cent a year, from $120 billion to $1.4 trillion. Over the same period, international bank lending rose from $324 billion to $7.5 trillion and the international bond market increased in size from $295 billion to $1.6 trillion. These sums dwarf the capital at the disposal of any national government or central bank.

Mobility of capital versus border controls on people

Whilst capital is free to roam the world, national governments deny its victims that right. In part, they maintain national barriers in order to better service the needs of transnational corporations, which require reliable supplies of skilled, as well as low-cost, labour. Governments

internationally are competing to supply cheaper, more trained and more disciplined labour forces to investors. At the same time, they also retain national restrictions on labour movement in an attempt to shore up their own sovereignty and domestic political control, using 'illegal' migrants as scapegoats for mounting social problems.

One law for the wealthy and another for the poor

Both overtly and covertly, national governments discriminate against the poor and working people when deciding whom to admit as migrants, temporary residents, visitors and students. Those who have wealth, particularly money to invest, or sought-after skills – which usually require means to acquire – are far more likely to be granted entry. In some cases, they can literally buy their way in.

Australia provides an example. Officially, 'Australia has a non-discriminatory immigration policy, which means that anyone from any country can apply to migrate'.[6] In reality, a potential immigrant's personal and/or family wealth is among the most significant factors in the assessment of their application. Some classes of visa are reserved specifically for applicants who can pass wealth or income tests. For instance, 'business skills' applicants obtain bonus points for having or bringing more money into the country. Thus, investors with $2 million obtain 80 of the 105 points needed for a visa.

The poor and most workers can only dream of these sums. They are further blocked by visa application fees that exceed more than a year's earnings for the average person living in a third world country, as well as charges for skills assessment, English language tuition and medical testing. Health requirements are another barrier. Poorer people have less access to medical care and healthier lifestyles, and some diseases, such as tuberculosis, are inherently associated with poverty.

Echoes of 'White Australia'

As numerous commentators have observed, Australia, partly because of its remoteness, faces only a trickle of refugees arriving on its shores compared to the countries of Western Europe and North America. Even if the figures are adjusted for population, Australia ranks 14th behind countries such as Switzerland, Belgium, Austria and the Netherlands.

Yet, the Australian government has been leading calls for harsher restrictions on asylum seekers, reviving memories of the 'White

Australia' policy that prevailed at the turn of the last century. One of the first pieces of legislation passed by the federal parliament was the *Immigration Restriction Act 1901*, directed at preventing the entry of non-whites. It was soon followed by the *Pacific Islanders Labourers Act*, which required the deportation of some 8 700 indentured South Pacific workers and their families. In the parliamentary debate, Labor Party MP and Australian Workers Union leader WG Spence summed up the program of the Labor and trade union leaders as follows:

> If we keep the race pure, and build up a national character, we shall become highly progressive people of whom the British government will be prouder the longer we live and the stronger we grow.[7]

The White Australia policy was only abandoned in the mid-1960s because of the growing dependence of Australian mining companies and graziers on Japan and other Asian markets. Right-wing populist politicians such as Pauline Hanson continue to advocate White Australia policies today. While the major parties – the Liberal-National Party Coalition and the Labor Party – formally eschew such policies, they maintain a considerable degree of bipartisan unity on the mandatory detention of asylum seekers, the restriction of detainees' legal rights and other measures designed to deter applicants for refugee status.

Under the banner of 'multiculturalism', they have sought to fashion a new national identity and international image but still within the framework of maintaining a relatively small population, insulated from the much larger populations of nearby Asia. This is a particular expression of a global divide. Despite the end of formal colonialism in most states, the world remains divided into oppressor and oppressed countries, with the advanced industrialised nations profiting at the expense of the so-called 'third world'.

To secure the means to finance investments, build infrastructure and operate basic facilities, the governments of the semi-colonial countries have to implement the 'structural adjustment' programs of the International Monetary Fund and the World Bank. These programs usually require draconian cuts to social programs, privatisation of state enterprises and tax concessions for international investors. Their effect is to transfer vast amounts of wealth into the coffers of the major banks and transnational corporations. In the words of one commentator:

> Not even at the height of its glory did the British Empire possess even a fraction of the power over its colonial subjects that the modern insti-

> tutions of world imperialism such as the World Bank, the IMF, GATT and the EC routinely exercise over the supposedly independent states of Latin America, Asia, Africa and the Middle East.[8]

Despite massive debt repayments, extracted at enormous social cost, the level of third world indebtedness continues to rise. In 1990, the total debt owed by developing countries was $1.4 trillion; by 1997, it had risen to $2.17 trillion. In 1998, third world countries paid $717 million in debt service to the major banks and financial institutions every day. These glaring disparities will continue to cause catastrophic conditions in many parts of Africa, the Middle East, Asia and Latin America, adding to the demand for means of entry to the advanced industrialised countries.

Defending rights within the Convention

Asylum seekers and those who oppose the treatment meted out to them will naturally seek to defend their rights to the fullest extent possible within the existing framework of the Convention. Unfortunately, the High Court of Australia has tended to restrictively interpret the scope of judicial review in the refugee context.

For example, there were strong grounds for the High Court to prevent the deportation of Kosovo and East Timorese asylum seekers, despite the provisions of the 2000 'safe haven' legislation, which denied them the right to apply for refugee status. However, in *Re The Minister for Immigration and Multicultural Affairs; Ex parte Fejzullahu* [2000] HCA 23, the court declined to hear the Kosovo refugees' cases.

Some commentators have argued that new uses can be found for the Convention in recognising international human rights claims on grounds of oppression related to gender and sexuality. Others have sought to find ways around the 'gaps' in the Refugee Convention by invoking other international instruments, including the Convention Against Torture and the International Covenant on Civil and Political Rights, to intervene on behalf of those needing protection.

Unfortunately, even in cases of extreme vulnerability, such as those asylum seekers being deported against their will, courts have failed to intercede, even though violent deportation methods can lead to death. In some cases, the individual actions of airline pilots or collective opposition, in the form of trade union bans, have saved deportees from disaster.

The Convention's fundamental flaws

More fundamentally, the fact remains that the Refugee Convention, even augmented by other treaties, does not assist the vast majority of the displaced persons. It is well established that the Convention is narrow and restrictive. As noted a decade ago by Hathaway:

> Most Third World refugees remain de facto excluded, as their flight is more often prompted by natural disaster, war, or broadly based political and economic turmoil than by 'persecution', at least as that term is understood in the Western context.[9]

The Convention is deficient in at least four primary respects. First, it does not protect the starving, the destitute, those fleeing war and civil war, or even natural disaster, let alone those seeking to escape economic oppression. Its narrow focus on *individuals* who are *persecuted* does not allow for mass exoduses in the face of suffering, injustice or discrimination that is not considered serious enough to amount to persecution.

Its requirement that this persecution be on the *specific* grounds of race, nationality, religious belief, political opinion or membership of a particular social group does not apply to people seeking refuge from torture, cruel punishment or other infringements of democratic rights, no matter how serious, despite efforts to extend the interpretation of 'particular social group' to include gender, sexual preference and child-bearing.

Second, the Convention does not create a right to enter another state, only a limited obligation on a national state not to expel or return a refugee to a state where they face persecution. In fact, the Convention does not recognise the individual's right to asylum; only the right of national states to decide who enters their territory. As stated by the High Court of Australia:

> The right of asylum is a right of States, not of the individual: no individual, including those seeking asylum, may assert a right to enter the territory of a State of which that individual is not a national.[10]

Third, even those accepted as refugees have no right to permanent residence and hence can be consigned to a tenuous and insecure status. The principle of *non-refoulement* under the Convention's Article 33(1) allows them to be removed to a so-called safe third country or to be forcibly repatriated to their home country once a government considers that the reasons for refugee status have ceased, as provided in Article 1C(5).

The Convention only assists asylum seekers who manage, invariably by means designated as 'illegal', to arrive physically in the country where they seek refuge. It does not impose any obligation on a country to take offshore applicants, that is, the overwhelming majority of people languishing in refugee camps throughout the poorest parts of the world, whether in their own countries or neighbouring states.

Governments refer to unwanted arrivals as 'queue jumpers', 'illegals' and 'forum shoppers'. Yet, refugees can only obtain the limited protection available under the Convention by escaping and entering a 'safe' country without permission. (Moreover, as stated earlier, the Convention upholds their right to do so, implicitly forbidding discrimination on the grounds of illegal entry.)

These fundamental flaws reflect the Convention's Cold War origins. It was drawn up in the aftermath of World War II and the Nazi Holocaust, which had caused the displacement of more than 40 million people within Europe. The knowledge that the advanced capitalist countries had refused to open their borders to many fleeing fascist persecution led to a broadly held sentiment that never again should refugees be turned away.

These democratic aspirations were incorporated in the Convention, which set out that all asylum-seekers – defined as those having a well-founded fear of persecution – were to be guaranteed certain inalienable rights, specifically that of refuge. Nevertheless, key governments only accepted the Convention on the basis that it did not create any duty to grant permanent residence and that they retained the sovereign right to decide which refugees were allowed entry to their countries.

Those who framed the Convention were also mindful of broader political considerations. In upholding the right to political asylum, the West sought to strengthen its democratic credentials against the Soviet Union and Eastern bloc countries, and specifically to hold the door open for political dissidents from the Stalinist regimes. The very conception of 'persecution' was tailored to give western governments ideological kudos for providing sanctuary to 'defectors' to the 'free world'.

Public opinion

Many writers in this field assert or assume a public opinion that is hostile to refugees and economic migrants. They tend to present governments as simply reacting to or appeasing this sentiment. Crock, for

example, after noting that the UN Human Rights Committee had condemned as 'arbitrary' (under the International Covenant on Civil and Political Rights) the automatic and indefinite imprisonment of unlawful entrants in Australian detention centres, wrote:

> If Australia does not respond in real terms to the Committee's rulings, the case may stand in this country as a testament to an increasing mood of national introspection and even isolationism from the world community.[11]

This supposed 'mood' is, however, one that is aggressively cultivated by those who hold political office and by the media proprietors. Crock herself noted instances of poor reporting and blatant scaremongering in the media, such as early 1998 reports that refugee claimants ('tourists') were coming to Australia for a '$30 work visa'.

Need for an alternative perspective

For public opinion on refugees to be genuinely tested, a viable alternative perspective must be advanced – one that corresponds to the requirements of global economic and social life and the needs and aspirations of the vast majority of people, rather than the vested interests of corporate and government elites.

A number of authors have suggested possible models for replacing the Convention with new international frameworks for protecting and assisting refugees. None of these models challenge the underlying assumption that nation-states and national borders will continue to exist throughout the 21st century. Instead, they seek ways to dilute the refugee obligations of nation-states according to what the authors consider politically palatable. Hathaway has argued specifically for tailoring proposals for change to meet the needs of national governments. 'In an international legal system based on the self-interest of states, it is critical that principled reform proceed in a manner which anticipates and responds to the needs of governments,' he stated, calling for support for a 'broader (if shallower) level of protection for most of the world's refugees'.[12]

In general, these authors invoke notions such as limited safe havens, temporary protection, international, regional and bilateral co-operation, and burden sharing. They also seek to separate the refugee regime from migration programs, suggesting that this will ease public concern over so-called asylum-driven migration and people smuggling.

This approach is based on maintaining the strict distinction between refugees and migrants. In a global world, and one increasingly dominated by social inequality, this is an artificial, inhumane and ultimately unreal perspective. The United Nations has estimated that 125 million people are, at any given time, outside their homeland in search of more secure political conditions or a better economic future. As a senior Canadian immigration official has observed:

> Almost all parts of the world are witnessing major migratory movements. While in 1965, 65 million people were living long term outside their countries of normal residence, by 1990 there were 130 million and in 2000 an estimated 150 million. Some are persons with legal status in their adopted countries. Most are in an irregular situation and try by various means to regularise their status.[13]

This demand for a more decent life will only grow amid ever wider disparities in wealth and life opportunities. According to the 1998 UN World Development Report, the three richest people in the world have assets exceeding the combined GDP of the 48 least developed countries, the 15 richest people have assets worth more than the total GDP of sub-Saharan Africa and the 32 richest more assets than the GDP of South Asia. Of the 4.4 billion people in so-called developing countries, almost three-fifths lack basic sanitation, one-third have no safe drinking water and one-quarter have inadequate housing, while one-fifth are undernourished, and the same proportion have no access to decent health services. According to the United Nations, out of the 147 countries defined as 'developing', some 100 had experienced 'serious economic decline' over the past 30 years.

Moreover, the advent of new forms of mass information, information technology and greater accessibility to air travel will accelerate and facilitate the movement of large numbers of oppressed people.

Citizenship and democratic rights

There is a profound connection between democratic rights and the rights of the most vulnerable in society: those denied entry to, or citizenship of, a country where they feel secure and able to participate meaningfully in political life.

Without the right to live securely with full political and social rights, democracy itself is meaningless. As one study noted, in the 20th century:

> The possession of a nationality became a matter of crucial, practical importance to the individual. The stateless person has no right of residence in any territory, no right to apply for employment or establish a business in any particular place. In some countries, a stateless person has only limited access to the legal system and its protection. One needs a nationality in order to enjoy basic security of residence somewhere.[14]

In Australia, in the 1992 case of *Lim*, the High Court upheld the power of the government to detain non-citizens ('aliens') indefinitely without trial, a practice that breaches one of the most fundamental democratic rights – freedom from arbitrary detention – which is enshrined in the centuries-old common law doctrine of habeas corpus. In the words of Brennan, Deane and Dawson JJ, Australian citizens enjoy a 'constitutional immunity from being imprisoned by Commonwealth authority except pursuant to an order by a court in the exercise of the judicial power of the Commonwealth'.[15]

However, the court found that this constitutional immunity did not extend to immigration detainees because they were not being incarcerated by way of criminal sanction but rather to protect the national interest. Remarkably, ignoring the reality of refugee persecution, the majority asserted that the detainees (Cambodian asylum seekers) were to some extent voluntary 'prisoners' who were free to return to their country of origin if they wished. In the final analysis, the High Court upheld the mandatory detention regime on the basis of the power to legislate with respect to aliens. The majority concluded that s 51 (xix) of the Australian Constitution conferred upon the executive authority to detain an alien for the purposes of expulsion or deportation, with an incidental power of detention to investigate and determine an application for asylum.

Several commentators have observed that the High Court failed to protect basic human rights, acting under pressure to accommodate to the vehement views of a government intent on having its way. Equally ominous was the readiness of the court to acknowledge that similar regimes of administrative detention could be applied to citizens during times of war, under the defence power. Once a precedent is established for the denial of basic democratic rights, first for the most vulnerable members of society – refugees and others without citizenship status – it can more easily be extended to others whose commitment to the state is called into question.

Global citizenship

The existence of nation-states, partitioning the globe into a patchwork of larger and smaller entities, each with their own border controls and exclusion regimes, is not natural or of ancient origin. Modern nationalism and general restrictions on the movement of people emerged only in the late 19th and early 20th centuries. Dummett and Nicol have pointed out that

> In earlier periods, restriction was by no means unknown but it was neither so general nor so systematic … Before the 1914 war, it was possible to travel between a number of countries without a passport, and with no restriction on taking work after arrival. With the price of a passage, an individual could take a free decision to look abroad for a new life; even without it, one could 'run away to sea', work a passage and try one country after another.[16]

From the late 19th century, however, border restrictions began to limit these movements.

> Instead of being 'chained to the soil' of a feudal lord, the twentieth-century poor gradually became chained to the territory of their countries of origin because other countries' rules forbade them entry.[17]

Far from nation-states being rooted in primordial nationalist sentiments, or even geographical necessity, their relatively recent historical origins are bound up with the socio-economic requirements of emerging capitalism in the struggle against the old feudal order in Europe. The growth of industrial production and the accumulation of capital required the development of a national market and the breaking down of guild privileges, political restrictions, local customs barriers and tariffs, which hemmed in production on all sides.

In England, the United States and France, revolutions were necessary to overthrow monarchical rule and establish nation-states. By the early 20th century, however, the enormous development of production engendered by capitalism had already outgrown the nation-state framework, leading to two world wars between the major economic and military rivals. Just as the old feudal fiefdoms, principalities and kingdoms had to be swept aside to clear the way for economic development under capitalism, it is now necessary to replace the nation-state system to allow for the harmonious growth of the world economy. That transformation is inseparable from establishing a new global conception of democracy and citizenship.

As currently instituted, citizenship is confined to a given nation-state, and does not extend beyond its borders. Meaningful democracy in the 21st century demands the right of all people to move wherever they wish around the world, the right to live, work and study wherever they choose, enjoying the political, civil and social rights and benefits available to all. If the poor and the oppressed are to be given the same right to travel and live as the wealthy, and if the right to immigrate as well as to emigrate is to be recognised, a new form of citizenship is needed: global citizenship.

Some commentators, while acknowledging the crisis of the nation-state system and its inability to deliver the democratic potential of globalised information technology, have dismissed this conception as utopian. Nevertheless, various attempts have been made to elaborate such a new paradigm. In an essay entitled 'The Making of Global Citizenship', Falk sought to combine traditional citizenship, still operating territorially, with global citizenship, operating temporally. He called for a sense of citizenship 'responsive to the varieties of human situation and diversity of cultural values', which would presuppose a reconstituting of the dominant political powers.[18]

In their book *Empire*,[19] Hardt and Negri identified new conceptions of identity and difference, networks of communication and control, and paths of migration, contending that they establish the basis for a truly democratic global society without national state borders. This is not the place to discuss the flaws in these analyses, but their emergence demonstrates an emerging recognition of the need for a global reshaping of human civilisation.

Discussion topics

1. Who is a refugee?
2. Assess current Australian policies toward asylum seekers.
3. Should the Refugee Convention be defended, or should a new approach be adopted?

Additional resources

S. Castles and M.J. Miller, *The Age of Migration: International Population Movement in the Modern World*, Macmillan, London, 1993.

Globalization and the International Working Class: A Marxist

Assessment, International Committee of the Fourth International, Mehring Books, Sydney, 1999.

M. Head, 'The High Court and the Removal of Kosovar Refugees' (2000) 4 *Macarthur Law Review* 197.

M. Head, 'The High Court and the Tampa Refugees: A Critical Examination of *Vadarlis v Minister for Immigration and Multicultural Affairs*' (2002) 11 *Griffith Law Review 23*.

M. Head, 'The Kosovar and Timorese 'Safe Haven' Refugees: A Test Case for Democratic Rights' (1999) 24 *Alternative Law Journal* 279.

M. Head, 'Review of Immigration and Refugee Law in Australia' [1999] *Australian International Law Journal 262*.

M. Head, 'When Fear of Death is not Sufficient for Refugee Status' (1998) 2 *Macarthur Law Review* 127.

E. Hobsbawm, *Nations and Nationalism since 1780*, Cambridge University Press, 1992.

The Immigration Kit, 6th edition, Federation Press, Sydney, 2001.

D. McMaster, *Asylum Seekers: Australia's Response to Refugees*, Melbourne University Press, Melbourne, 2001.

P. Mares, *Borderline: Australia's Treatment of Refugees and Asylum Seekers*, UNSW Press, Sydney, 2001.

The Problem with the 1951 Refugee Convention, Parliamentary Library, Research Paper No 5, 2000–01, Parliament of Australia, Canberra.

NOTES

INTRODUCTION

1 D. Pearce, E. Campbell, E. Harding and D. Harding and Commonwealth Tertiary Education Commission (1987), *Australian Law Schools: A Discipline Assessment for the CTEC*, 1987.
2 C. McInnis and S. Marginson, *Australian Law Schools after the 1987 Pearce Report*, 1994, p 157.

CHAPTER 1

* The inspiration for the overall structure and much of the content of this chapter and chapter 2 is owed to Patrick J. Hurley's *Concise Introduction to Logic*, 5th edition, Wadsworth Publishing Co, Belmont California, 1994.
1 Hurley, p 1.
2 Ibid.
3 Ibid, p 2.
4 Ibid, p 6.
5 A. MacAdam and J. Pyke, *Judicial Reasoning and the Doctrine of Precedent in Australia*, Butterworths, Sydney, 1989, p 41, quoting Cross.
6 [1967] 1 QB 443 at 502.
7 L. Waller, *Derham, Maher and Waller: An Introduction to Law*, 8th edition, LBC Information Services, Sydney, 2000.
8 R. Pine, *Essential Logic*, Harcourt Brace, Fort Worth, 1996, p 43.
9 Hurley, pp 14–21.
10 Ibid, p 19.
11 J.G. Fleming, *The Law of Torts*, 8th edition, Lawbook Co, Sydney 1992, p 194.
12 Pine, pp 233–4.
13 Hurley, pp 21–4.
14 J. Burbidge, *Within Reason,* Broadview Press, Peterborough, 1990, pp 99–100.
15 Ibid, p 79.
16 Hurley, p 31.
17 See for example C. Cook, R. Creyke, R. Geddes and I. Holloway, *Laying Down the Law*, Butterworths, Sydney, 2001, p 58.
18 Hurley, p 41.

19 Ibid, p 42.
20 There are different levels of logical structure. At one level the arguments below have the invalid structure: A. B. Therefore C. This is why we must specify the 'most specific' level of structure in considering validity: that is, taking account of all 'logical words'.
21 A. Fisher, *The Logic of Real Arguments*, Cambridge University Press, 1988, p 142.
22 Ibid.
23 It is possible that premises and conclusion are true, depending upon the particular content involved, as, for example, where A = you are human, B = you are warm-blooded and C = you are a mammal. But the argument remains invalid.
24 Hurley, pp 44–5.
25 Ibid, p 47.
26 Ibid, p 34.
27 Ibid, p 35.
28 Ibid, p 468.
29 Burbidge, p 18.
30 Ibid, pp 18–19.
31 Hurley, p 524.
32 Ibid, p 468.
33 Ibid, p 64.
34 J. Nolt and D. Rohatyn, *Logic*, McGraw Hill, New York, 1988, p 18.
35 I. Campbell, *Sydney Morning Herald*, 29 March 2001.

CHAPTER 2

1 Of course, while the factual issues might be straightforward in theory, it is not necessarily easy to provide definitive answers to such questions in practice. Chapter 4 considers some of the difficulties. And it could well be that the death penalty – or its abolition – has very significant unrecognised consequences that could profoundly influence ethical decision-making if they were recognised.
2 See for example S. George, *Another World Is Possible if...*, Verso, London, 2004, p 79.
3 R. Murugason and L. McNamara, *Outline of Criminal Law*, Butterworths, Sydney 1997, pp 170.
4 K. Hall, *Legislation*, Butterworths, Sydney, 2002, pp 73–4.
5 Ibid, p 78.
6 D. Khoury and Y. Yamouri, *Understanding Contract Law*, Butterworths, Sydney, 1995, p 89.
7 MacAdam and Pyke, p 267.
8 Ibid.
9 Ibid.
10 Ibid, pp 268–9.
11 Ibid, p 40.
12 Ibid.
13 Ibid, p 278.
14 Ibid, p 279.
15 See chapter 3.
16 See J. Conaghan and W. Mansell, *The Wrongs of Tort*, Pluto Press, London, 1993, pp 70–84.
17 Ibid, p 73.

CHAPTER 3

1 See R. Nader and W. Smith, *No Contest*, Random House, New York, 1996, pp 14–15.
2 Ibid, p 15.
3 Ibid.

4 Ibid, p 16.
5 Ibid, p 17.
6 Ibid, pp 17–18.
7 This example is drawn from a recent law textbook: P. Keyzer, *Legal Problem Solving*, Butterworths, Sydney, 2003, p 19. Unfortunately, Keyzer confuses validity with soundness, misrepresents the structure of *modus ponens* (confusing a conditional 'then' with a 'then' representing temporal sequence) and misrepresents *modus tollens* (confusing it with denying the antecedent).
8 P.J. Hurley pp 116–120.
9 Ibid, p 124.
10 S. McNicol and D. Mortimer, *Evidence*, Butterworths, Sydney, 1996, pp 103–4.
11 Hurley, p 127.
12 G. Gigerenzer, *Reckoning with Risk*, Penguin, London, 2002, pp 142–3. See also Appendix on probability.
13 Ibid, p 143.
14 Ibid, p 144.
15 Ibid, pp 144–5.
16 Ibid, p 154.
17 Gigerenzer, p 165.
18 Ibid. Claims of '1 in 10 million people with DNA base sequence X' actually derive from the multiplied frequencies of sub-sequences in many smaller sample groups. But there might be much less variability in particular – isolated – groups.
19 Ibid.
20 They also devised the playing card 'test' referred to earlier.
21 B.S. Everitt, *Chance Rules*, Springer-Verlag, New York, 1998, p 100.
22 Ibid, p 101.
23 Ibid, pp 108–9. In this case, the appropriate formula is as follows:
    1 Pr(innocent/DNA match) = Pr(DNA match/innocent) (= RMP) x Pr(innocent)/Pr(match)
    Here, Pr(match), the unconditional probability of a match, is the difficult one. Here is where we need to take account of the population of possible perpetrators (for example, 1 million people).
    2 Pr(match) = Pr(match/innocent) x Pr(innocent) + Pr(match/guilty) x Pr(guilty); thus if
    3 Pr(match/innocent) = 0.000001 (1 in a million randomly sampled folk would match the sample); and
    4 Pr(innocent) = 0.999999 (out of 1 million possible perpetrators, only 1 actually did it); then
    5 Pr(guilty) = 0.000001 [1/M]; and
    6 Pr(innocent/match) = 0.000001 (Pr match/innocent) x 0.999999 (Pr innocent)/[(0.000001 x 0.999999) + (1 x 0.000001)] (Pr match) = 0.5 (approx.)
24 Moving on to the DNA case, we consider 1 000 000 men who could be responsible for the crime. 1 actually did it. 999 999 didn't do it. Assuming no false negatives, a DNA test given to that 1 will be positive. If DNA tests are given to all the other 999 999, 1 will be positive [1/M(RMP) x 999 999/1], the rest will be negative (assuming no false positives). So Pr(guilt/match) = 1/2.

CHAPTER 4

1 Hurley, pp 143–4.
2 J. Swanton, B. McDonald and R. Anderson, *Cases on Torts*, Federation Press, Sydney, 1992, p 136.
3 Ibid.
4 Ibid.
5 Ibid, p 137.

6 It is true that serial killers are frequently said to target their victims in random and unpredictable ways, often over large areas. But this does not mean that diligent investigation cannot effectively narrow the focus of investigation.
7 Hurley, p 138.
8 See for example A. Farrell, *Crime, Class and Corruption*, Bookmarks, London, 1992.
9 Hurley, p 139.
10 Ibid, p 142.
11 Cook et al, p 103.
12 MacAdam and Pyke, p 297.
13 Ibid, p 291.
14 Swanton et al, p 137.
15 Ibid.
16 In England in 1977 and 1978 the 'disappearance' of 28 sackfuls of captured cannabis and the adulteration of another major police drug haul with fingerprint powder received substantial publicity. See Farrell, p 40. The drug squad at Scotland Yard was disbanded in the early 1970s due to corruption and the Obscene Publications Squad was identified by Justice Jones as a 'vast protection racket' in 1977. Eighty-two officers were sacked and another 301 left during criminal or disciplinary enquiries between 1973 and 1976 under reforming Commissioner Robert Mark. By 1978 Operation Countryman had compiled evidence against 78 officers from Scotland Yard and 189 from the City of London police involved in armed robberies. In the early 1980s the Chief Constable of Cheshire was looking into allegations of misconduct on Humberside. Humberside was conducting investigations into South Yorkshire drug squad and this investigation itself was later looked at by Merseyside Police.

    In 1989, the framing of innocent people, fabrication of evidence and lies on oath leading to the wrongful imprisonment of the Guildford Four were widely discussed in Britain. And the whole of the West Midlands Regional Crime Squad was disbanded amid accusations of forced confessions and tampering with evidence. That year there were 37 deaths in police custody in the Metropolitan Police area alone. As Farrell points out, 'Only 3 out of 17 coroners' inquests into such cases in 1989 found verdicts of "natural causes".' [And] there were 2372 complaints of assault carried out by the London police in 1989'. Farrell, p 19.
17 In Canada, in the case of *Jane Doe v Metropolitan Toronto [Municipality] Commissioners of Police* [1998] 160 DLR (4th) 697, the plaintiff was successful in establishing that the MTPF had conducted a negligent investigation and failed to warn women they knew to be potential targets of a rapist. '[A]s a result of such conduct, [the rapist] was not apprehended as early as he might have been and the plaintiff was denied the opportunity … to take any specific measures to protect herself.' Similarly, in the German case of BGM LM S839[fg] BGB Nr 5, in 1953 the plaintiff successfully sued a police organisation for breach of official duty, involving failure to arrest known criminals and thereby prevent criminal acts. But this still contrasts starkly with English cases that have followed the Hill line even when the identity of the offender is known to the police. And as Surma notes:

    > the reasoning of the German court in its entirety suggests that … there is an official duty of the police to prevent criminal acts owed to possibly affected third parties whenever the police are in an obvious position to prevent serious crimes'. R. Surma, 'A Comparative Study of the English and German Judicial Approach to the Liability of Public Bodies in Negligence' (2000) *Oxford University Comparative Law Forum* 8.

18 A classic example here is the drug diethylstilbestrol widely prescribed to pregnant women to prevent miscarriage from 1940 to 1971.
19 R. Fogelin and W. Sinnot-Armstrong, *Understanding Arguments*, 5th edition, Harcourt Brace, Fort Worth, p 340.
20 McNicol and Mortimer, p 287.
21 Ibid, p 287.

22 Ibid, p 288.
23 We can here refer to the case of Fred Zain, a state serologist, who lied and systematically manufactured evidence in at least 133 rape and murder cases in West Virginia, leading to convictions of numerous innocent people. See D. Walton, *Appeal to Expert Opinion*, Pennsylvania State University Press, 1997, p 181.
24 Hurley, p 152.
25 J. Smith, *Misogynies*: *Reflections on Myth and Malice*, Faber and Faber, London, 1989, pp 118, 121–22.
26 Ibid, pp 123–4.
27 In 1998 the European Court of Human Rights (ECHR) rejected a decision by the British Court of Appeal in *Osman v Ferguson*, in which the Court of Appeal had relied upon Hill to reaffirm police immunity from liability in negligence in respect of their activities in the investigation and suppression of crime. The ECHR held that this decision was a violation of the plaintiff's civil rights to a fair and public hearing of her complaints. Despite this, Hill has continued to stand as 'good law' protecting the police in England and in Australia from any liability in negligence. On 29 March 2004, Hill was cited in *Cran v State of New South Wales* (2004) NSWCA 92. The appellant claimed to have suffered harm in gaol by way of chronic post traumatic stress disorder because the Police and the Director of Public Prosecutions neglected to secure prompt analysis of suspected prohibited drugs found in his possession. The Appeal Court supported an earlier judgment that since the omission in question 'occurred in course of police investigation, it was immune from civil action.
28 Hurley, p 155.
29 McNicol and Mortimer, p 88.
30 Ibid, p 87.
31 Ibid, p 95.
32 D. Brown, D. Farrier, S. Egger and L. McNamara, *Criminal Laws*, Federation Press, Sydney, 2001, p 232.
33 Hurley, p 157.

CHAPTER 5

* This chapter draws significantly upon Ronald Giere's *Understanding Scientific Reasoning*, 1st (1979), 2nd (1984) and 4th (1997) editions respectively, Holt, Rinehart and Winston, Fort Worth.
1 T. Schick, Jr and L. Vaughn, *How To Think About Weird Things*, Mayfield, Mountain View, 1999, p 151.
2 Ibid.
3 A. Chalmers, *What Is This Thing Called Science?*, 2nd edition, Queensland University Press, St Lucia, 1982, pp 2–5.
4 Shick and Vaughn, p 162.
5 This terminology is owed to philosopher of science Rom Harre.
6 Isolation and control generally cannot be achieved in relation to large-scale social phenomena. But sampling and surveying are still relevant here, including cross-cultural comparisons.
7 Giere, 2nd edition, p 177.
8 Ibid, p 180.
9 R. Hogg, in D. Brown and M. Wilkie (eds), *Prisoners as Citizens*, Federation Press, Sydney, 2002, p 11.
10 Ibid, p 14.
11 Ibid, p 15.
12 Ibid, p 10.
13 Here we are using a streamlined statistical procedure, and concrete examples, drawn from Ronald Giere's textbook on scientific reasoning, *Understanding Scientific Reasoning*, 4th edition, Harcourt Brace, Orlando, 1997.

14 Ibid, p 160.
15 J. Philips, *How To Think about Statistics*, 6th edition, W.H. Freeman and Co, New York, 2000, pp 150–1, p 193.
16 D. Rowntree, *Statistics without Tears*, Penguin, London, 1981, pp 156–7.
17 J.T. Dennis, *The Complete Idiot's Guide to Physics*, Alpha Books, Indianapolis, 2003, p 51.
18 Ibid, p 52.
19 L. Schultz, 'US Foreign Policy and Human Rights Violations in Latin America: A Comparative Analysis of Foreign Aid Distributions', *Comparative Politics*, Vol 13, No 2, January 1981, p 155; M. Stohl, D. Carleton and S. Johnson, 'Human Rights and US Foreign Assistance from Nixon to Carter', *Journal of Peace Research*, Vol 21, No 33, 1984, p 217; D. Carleton and M. Stohl, 'The Foreign Policy of Human Rights: Rhetoric and Reality from Jimmy Carter to Ronald Reagan', *Human Rights Quarterly*, Vol 7, 2, May 1985, p 215.
20 F. Gareau, *State Terrorism and the United States*, Zed Books, London, 2004, pp 220–1.
21 Ibid.
22 Ibid.
23 Ibid.

CHAPTER 6

1 Fleming, p 193.
2 Ibid, pp 193–4.
3 Ibid, p 203.
4 R. Martin, *Scientific Thinking*, Broadview Press, Peterborough, 1997, chs 18 and 19.
5 Ibid.
6 Ibid, pp 275–6.
7 Ibid.
8 Ibid.
9 Informed of the results of all previous tests; whether this drug is just a variation on a competitor's product with no real health benefits; provision of research funding by taxpayers; availability of the drug to those that need it.
10 Martin, chs 18, 19.
11 Giere, 4th edition, p 229.
12 Ibid.
13 Ibid, p 241. Given this sort of information, available for decades, obvious questions arise about the failure of criminal and civil law.
14 Ibid.
15 Ibid, p 233.
16 Ibid, p 237.
17 Ibid, p 238.
18 Hurley, p 524.
19 *New Scientist*, 6 March 2004, p 19.
20 R. Moynihan, *Too Much Medicine?*, ABC Books, Sydney, 1998, pp 54–5.
21 Ibid.
22 Gigerenzer, p 34.
23 Ibid.
24 Ibid, p 35.
25 D. Gee and M. Greenberg in P. Harremoes, D. Gee, M. McGarvin, A. Stirling, J. Keys, B. Wynne and S. Guedes Vaz, *The Precautionary Principle in the 20th Century*, Earthscan, London, 2002, pp 50–1.
26 Ibid, p 58.
27 A. Deville and R. Harding, *Applying the Precautionary Principle*, Federation Press, Sydney, 1997, p 19.
28 Harroemoes et al, p 51.

29 Ibid, pp 13–14.
30 Ibid.
31 Harremoes et al, p 191.
32 Deville and Harding, pp 71, 73.
33 Harremoes et al, p 189.
34 Ibid.
35 Ibid.
36 Ibid, p 192.
37 R.G. Wilkinson, *Unhealthy Societies*, Routledge, London, 1996, p 156.
38 Ibid, p 157.
39 E. Handsley, 'Market Share Liability and the Nature of Causation in Tort' (1994) *Torts Law Journal* 24 at pp 24–44.

CHAPTER 7

1 Schick and Vaughn, p 161.
2 See chapter 1.
3 Giere, 2nd edition, p 151.
4 Giere, 1st edition, p 137.
5 R. Norman, *On Humanism*, Routledge, London, 2004, p 43–4.
6 We see here a connection with some types of natural law theory discussed in later chapters. A womb is for bearing children, so abortion in unnatural and should be illegal.
7 Norman, p 36.
8 Ibid, p 37.
9 E. Mayr, *One Long Argument*, Penguin, 1992, p 22.
10 Norman, p 36.
11 Ibid, p 38.
12 D. Walton, *Appeal to Expert Testimony*, Pennsylvania State University Press, 1997, pp 182.
13 Ibid.
14 Ibid, p 184.
15 S. Odgers, *Uniform Evidence Law*, Federation Press, Sydney, 1995, pp 126–9.

CHAPTER 8

* This chapter draws significantly upon Wilfred J. Waluchow's recent book, *The Dimensions of Ethics*, Broadview Press, Peterborough, Ontario, 2003. So too does it draw upon the work of moral philosopher Richard Norman.
1 See Waluchow, pp 35–6.
2 Ibid, pp 36–7.
3 Ibid, pp 40–2.
4 See for example R. Harre and M. Krausz, *Varieties of Relativism*, Blackwell, Oxford, 1996, ch 5.
5 J.R. Searle, *Rationality in Action*, MIT Press, Cambridge, 2001, p 224.
6 Ibid, pp 232–3.
7 See S. Freud, *The Standard Edition of the Complete Psychological Writings of Sigmund Freud*, J Strachey et al (ed), 24 vols, Hogarth Press, London, 1953–74, 1930; *Civilization and its Discontents*, in *Standard Edition*, vol 21, pp 59–145.
8 See Waluchow, pp 42–5.
9 Ibid, pp 45–50. As Waluchow points out, the term 'right' is also sometimes used to refer to the power to alter existing normative relationships, as where I grant my lawyer the right to act upon my behalf in various legal transactions. Privileges are created when rights are voluntarily waived or suspended thereby relieving others of the corresponding obligations. Ibid, pp 50–3.

10 Ibid, pp 47–8.
11 False hopes in this regard have pernicious social consequences in encouraging support for a status quo built upon systematic obfuscation.
12 Civil rights have become increasingly 'formal' and 'empty' rights for most of the population of the English-speaking world as increasing costs of effective legal and political action have put these things far beyond the reach of the great majority of people.
13 R. Chisholm and G. Nettheim, *Understanding Law*, 6th edition, Butterworths, Sydney, 2002, p 126.
14 Tax laws do not currently ensure genuine progressive taxation and do not deal with the disposition of tax receipts.
15 This idea is taken seriously in the Scandinavian countries. Recently one of Finland's richest men was fined $283 419 for speeding through the centre of the capital. Finnish traffic fines vary according to the offender's income.
16 H. Glasbeek, *Wealth by Stealth*, Between the Lines, Toronto, 2002, p 154.
17 Ibid, p 158.
18 Ibid.
19 See Waluchow, pp 58–9.
20 See for example R. Norman, *The Moral Philosophers*, Clarendon Press, Oxford, 1983, ch 2.
21 R. Norman, *Ethics, Killing and War*, Cambridge University Press, 1995, p 7.
22 Ibid, p 8.
23 Ibid, p 10.
24 Waluchow, pp 96–7.
25 Ibid.
26 Ibid, p 101.
27 Ibid, p 102.
28 Ibid, pp 67–8.
29 Ibid, ch 3.
30 P. Singer, *One World*, Text Publishing, Melbourne, 2002, pp 125–6.
31 Ibid, p 126.
32 Ibid, p 131.
33 The 1998 Rome Statute defines genocide as 'any of the following acts committed with intent to destroy, in whole or in part, a national, racial or religious group, as such: (a) killing members of the group (b) causing serious bodily or mental harm to members (c) deliberately inflicting on the group, conditions of life calculated to bring about its physical destruction in whole or in part (d) imposing measures intended to prevent births within the group (e) forcibly transferring children of the group to another group. Ibid, pp 137–8.
34 'Crime against humanity' means any of the following acts when committed as part of a widespread or systematic attack directed against any civilian population, with knowledge of the attack: (a) murder (b) extermination (c) enslavement (d) deportation or forcible transfer of population (e) imprisonment or other severe deprivation of physical liberty in violation of fundamental rules of international law (f) torture (g) rape, sexual slavery, enforced prostitution, forced pregnancy, enforced sterilisation, or any form of sexual violence of comparable gravity (h) persecution against any identifiable group or collectivity on political, racial, national, ethnic, cultural, religious, gender … or other grounds that universally recognised as impermissible under international law, in connection with any act referred to in this paragraph or any crime within the jurisdiction of the Court (i) enforced disappearance of persons (j) the crime of apartheid (k) other inhumane acts of a similar character intentionally causing great suffering, or serious injury to body or to mental or physical health. Ibid, pp 138–9.
35 According to Article 8 of the ICC statute, war crimes include wilful killing, torture, wilfully causing great suffering or serious injury to body or health, extensive destruction of property not justified by military necessity, forcing prisoners to serve in the

armed forces of the other side, depriving prisoners or civilians of a fair trial, unlawful deportation and taking of hostages. Article 8.2(b) lists 26 other indictable crimes including offences against civilians and civilian targets, various sexual crimes, conscripting children into armed forces and starvation of civilian populations. See for example R. Piotrowicz and S. Kayr, *Human Rights in International and Australian Law*, Butterworths, Sydney, 2000, ch 9.

36 Singer, p 139.
37 Ibid, pp 155–6.
38 Gareau, p 223.
39 Ibid, p 227.

CHAPTER 9

1 Waluchow, pp 146–7.
2 Ibid, p 152.
3 Ibid, p 167.
4 Norman, *On Humanism*, p 102.
5 Ibid, p 104.
6 Such interventions could be costly and ineffective in reducing the harmful behaviour in question; or the harm caused by enforcement of relevant penalties, particularly imprisonment, could outweigh the harm of the original behaviour, etc.
7 I. Kant, *Groundwork of the Metaphysics of Morals*, H.J. Paton (trans), Harper and Row, New York, 1964, p 70.
8 Norman, *The Moral Philosophers*, p 52.
9 Ibid.
10 Ibid, p 54.
11 Waluchow, p 233.
12 Ibid, p 232.
13 Ibid, p 233.
14 Ibid.
15 C. Gilligan, *In a Different Voice: Psychological Theory and Women's Development*, Harvard University Press, Cambridge, 1982.
16 Nader and Smith, pp 334–5.

CHAPTER 10

1 M. Freeman, *Lloyd's Introduction to Jurisprudence*, 6th edition, London, Sweet & Maxwell, 1994, p 90. *Lloyd's* (chs 3, 4 and 6) is the main reference for this chapter, including the passages cited from legal theorists. The other main sources are M. Davies, *Asking the Law Question*; H. McCoubrey and N. White, *Textbook on Jurisprudence*, 3rd edition, London, Blackstone Press, 1999, chs 2, 3, 4 and 5; R. Wacks, *Swot Jurisprudence*, 5th edition, Blackstone Press, London 1999, chs 3, 4 and 5.
2 McCoubrey and White, p 12.
3 Woluchow, p 107.
4 Ibid, pp 104–16.
5 *Lloyd's*, p 109.
6 Ibid, p 137.
7 Ibid, pp 114–15.
8 C.B. Macpherson, *The Political Theory of Possessive Individualism: Hobbes to Locke*, Oxford, Clarendon Press, 1962, pp 199–200.
9 (1765) 19 How St Tr 1029.
10 *Lloyd's*, pp 116–17.
11 McCoubrey and White, p 81.
12 D. North, *Equality, the Rights of Man and the Birth of Socialism*, Detroit, Mehring Books, 1996, pp 10–11.

13 *Lloyd's*, pp 120–1.
14 *Lloyd's*, p 28.
15 *Lloyd's*, pp 122–3.
16 Ibid, p 250.
17 *Lloyd's*, p 224.
18 McCoubrey and White, p 26.
19 Ibid, p 32.
20 *Lloyd's*, p 343.
21 McCoubrey and White, pp 87–8.
22 *Lloyd's*, pp 157–61.
23 Ibid, p 171.
24 *Lloyd's*, pp 173 ff; McCoubrey and White, pp 96–7.
25 For a biting critique of Finnis, see Davies, pp 67–72.
26 Waluchow, pp 135–8.
27 Ibid.
28 According to Waluchow's model, everyone has an initial bargaining position and wishes to improve their position. Co-operation creates a surplus (of good things, over and above what could be created without it). Different agreements distribute it in different ways. The way proposed in this case is based upon consideration of the proportion that the absolute magnitude of concessions bear to the difference between the utility the person would derive from a particular decision and their utility in the initial bargaining position. Person H chooses action A, which will get them from 0 to 10 units of utility, and person S from 0 to 15 units. Person S chooses action B, which will get them from 0 to 30 units and H from 0 to 3 units. H's concession in Action A is 0, S's concession is 50% (15/30); S's concession in action B is 0, H's concession is 70% (7/10). Therefore action A is preferable.
29 See *Lloyd's*, pp 523–4.
30 Ibid, p 525.
31 Ibid, p 529.
32 Ibid, pp 590–3.
33 Ibid.
34 Ibid, pp 557–63, 647–57.
35 Ibid, 6th edition, p 462.
36 R. Posner, *The Economic Analysis of Law*, 2nd edition, Little, Brown & Co, Boston, 1977, p 202.

CHAPTER 11

1 R. Swift, *The No-Nonsense Guide to Democracy*, Verso, London, 2002, p 46.
2 Ibid, p 21.
3 A. Levine, *Arguing for Socialism*, Routledge, London, 1984, p 29.
4 Not all 'right liberals' have endorsed this idea of 'just deserts'. But it is certainly part of the 'popular ideology' of neo-liberalism.
5 P. Devine, *Democracy and Economic Planning*, Cambridge University Press, 1988, pp 109–10.
6 A. Callinicos, *An Anti-Capitalist Manifesto*, Polity Press, London, 2003, p 126.
7 Swift, p 105.

CHAPTER 12

1 F. Fukuyama, *The End of History and the Last Man,* London, Penguin, 1992.
2 For example, A. Hunt, *Explorations in Law and Society*, London, Routledge, 1993.
3 Pashukanis' main treatise, *Law and Marxism: A General Theory*, was originally published in the Soviet Union in 1924. It was published in a new English translation in 1978 (Inks Links, London).

4 K. Marx, Preface to *A Contribution to the Critique of Political Economy*, in K. Marx and F. Engels, *Selected Works*, vol. 1, Moscow, Progress Publishers, 1969, pp 503–4.
5 K. Marx, *The Civil War in France*, Moscow, Progress Publishers, 1948, p 17.
6 K. Marx, *Critique of the Gotha Program*, New York, International Publishers, 1970, p 10.
7 L. Trotsky, *The Revolution Betrayed. What is the Soviet Union and Where is it Going?*, New York, Pathfinder Press, 1972, pp 39–40. See also L. Trotsky, *Terrorism and Communism*, New Park Publications, London 1975, ch 3, 'Democracy'.
8 F. Engels, Preface to Marx, *The Civil War in France*, p 16.
9 F. Engels, Letter to Conrad Schmidt, 27 October 1890, in K. Marx and F. Engels, *Selected Correspondence*, Progress Publishers, Moscow, 1975, p 402.
10 F. Engels, Letter to J. Bloch, 21 September 1890, in ibid, pp 394–5.
11 Engels, Letter to Conrad Schmidt, in ibid, pp 399–402.
12 Ibid, p 482.
13 F. Engels, Letter to Mehring, 14 July 1893, in ibid.
14 K. Marx, *Capital*, vol 1, Lawrence & Wishart, London, 1970, p 80.
15 Trotsky, *Terrorism and Communism*, pp 60–1.
16 Ibid, p 9.

CHAPTER 13

1 M.D.A. Freeman, *Lloyd's Introduction to Jurisprudence*, Sweet & Maxwell, London, 2001, p 557.
2 R. Hahnel, *The ABC's of Political Economy*, Pluto Press, London, 2002, pp 31–2.
3 Ibid, p 77.
4 F. Stilwell, *Changing Track*, Pluto Press, Sydney, 2000, p 282.
5 Ibid, p 283.
6 D. Korten, *When Corporations Rule the World*, Earthscan Publications, London, 1997, p 79.
7 Stilwell, p 281.
8 Ibid, p 282.
9 P. Self, *Rolling Back the Market*, Macmillan, London, 2000, pp 27–8.
10 Ibid.
11 Hahnel, p 83.
12 Ibid, p 85.
13 Ibid, p 86.
14 Ibid, p 87.
15 Ibid, p 93.
16 Ibid.
17 Ibid.
18 Ibid, p 94.
19 Ibid, p 93.
20 Ibid, p 83.
21 S. Kangas, *Ronald Coase and the Coase Theorem*, p 3, <http://www.huppi.com/kangaroo/L-chicoase.htm>.
22 Ibid.
23 Ibid.
24 J. Bellamy Foster, *The Vulnerable Planet*, New York, 1999, p 132; A. Callinicos, *An Anti-Capitalist Manifesto*, Polity, London, 2003, p 111.
25 Kangas, p 4.
26 Ibid.
27 C. Harman, *Explaining The Crisis*, Bookmarks, London, 1984, p 17.
28 'Nick Beams Replies to a Reader's Question about the Law of the Falling Rate of Profit', World Socialist Web Site, 29 March 2000, <http://www.wsws.org/articles/2000/ mar2000/nb-m29.shtml>.

29 K. Marx, *The Grundrisse*, Random House, New York, 1973, p 340.
30 N. Beams, *The Significance and Implications of Globalisation*, Mehring Books, Sydney, 1998.

CHAPTER 14

1 See Glasbeek, ch 2.
2 M. Kelly, *The Divine Right of Capital*, Berret-Koehler Publishers Inc, San Francisco, 2003, pp 33–4.
3 K. Singh, *Taming Global Financial Flows*, Zed Books, London, 2000, p 19.
4 Glasbeek, ch 4.
5 J. Bakan, *The Corporation*, Free Press, New York, 2004, p 37.
6 Glasbeek, pp 325–6
7 Ibid, p 76.
8 Ibid, p 77.
9 Bakan, p 64.
10 Ibid, p 60.
11 Ibid.
12 Ibid, pp 62–3.
13 Ibid, p 63.
14 Ibid, p 64.
15 Ibid, p 75
16 Ibid, pp 75–9.
17 Ibid, p 79.
18 H. Shutt, *A New Democracy*, Zed Books, London, 2001, pp 48–9.
19 <http://www.accc.gov.au/content/index.phtml/itemId/259608/fromItemId/6106>, p 1.
20 <http://www.accc.gov.au/content/index.phtml/itemId/496100/fromItemId/142>, p 1.
21 In practice, collusion in fixing prices can be tacit, or passive collusion based upon mutual understanding, while the law requires evidence of active collaboration.
22 As Susan George points out, a firm producing raw materials in one country sells such materials to a manufacturing affiliate in another country and the finished product is marketed by a sales office of the same corporate group in a third country.
23 F. Stillwell, *Political Economy*, Oxford University Press, Melbourne, 2002, pp 180, 182.
24 T. Cleaver, *Economics: The Basics*, Routledge, London, 2004, p 61.
25 Ibid, p 74.
26 Shutt, p 51.
27 Stillwell, p 186.
28 S. Beder, *Power Play*, Scribe Publications, Carlton North, 2003, p 19.
29 Shutt, p 51.

CHAPTER 15

1 See for example Milton Friedman and Rose Friedman, *Free to Choose*, Penguin, 1980. Not all of the libertarian founding fathers of neo-liberalism have subscribed to these sorts of ideas. Frederick von Hayek, in particular, specifically rejected them. However, neoclassical economic influences have generally been much stronger than those of the Austrian School, and neoclassical theory points significantly towards 'just deserts'.
2 In their classic libertarian study cited above, Milton and Rose Friedman defend both remuneration in proportion to 'the price a person receives for the services of his resources' and in proportion to how 'hard' they work. Ibid, p 43.
3 H. Shutt, *A New Democracy*, Zed Books, London, 2001, pp 70–1.
4 I. Watson, J. Buchanan, I. Campbell and C. Briggs, *Fragmented Futures*, Federation Press, Sydney, 2003, p 23.
5 Ibid, p 39.
6 See for example <http//bosswatch.labor.au/>.

7 B. Ehrenreich, *Nickel and Dimed*, Granta, London, 2002.
8 Self, p 27.
9 See for example Callinicos, ch 3.
10 See for example R. Nozick, *Anarchy, State and Utopia*, Blackwell, 1974.
11 See for example P. Singer and T. Trainer, *The Greens*, Text Publishing, Sydney, 1996. In practice, monopoly mark-ups inflate prices while savings on externalities bring them down compared to the price of genuine sustainability. The funds that should support such sustainability go into the pockets of shareholders and executives.
12 Hahnel, p 97.
13 See for example M. Head, 'Whither the Refugee Convention? A New Perspective for the 21st Century', *Mots Pluriels*, No 21, May 2002.
14 See for example M. Rowbotham, *Goodbye America*, John Carpenter, Charlbury, 2000, ch 5.
15 Ibid.
16 Hahnel, ch 8.
17 Ehrenreich, p xi.
18 Hahnel, pp 282–4.

CHAPTER 16

1 Friedman and Friedman, pp 27–43.
2 Death taxes, or estate taxes or duties, have existed – in various, more or less limited – forms, around the developed world, for a long time. But the trend is towards their abolition rather than their extension, often on grounds of the 'unfairness' of double taxation of assets. In fact, there is no double taxation in this case, since dead people don't pay taxes, while working people's income is taxed at source and then again and again through consumption and other taxes. Similarly, taxing undistributed corporate profits and dividends is only double taxing of assets if we see shareholders as sole legitimate owners of all company profit.
3 Kelly, p 2.
4 Ibid, pp 33–5.
5 Ibid, p 45.
6 Ibid.
7 Ibid, p 46.
8 Ibid, p 2.
9 Ibid, p 6.
10 Ibid, p 60.
11 Ibid, p 51.
12 G. Hague, *Quarterly Essay*, issue 10, p 5; <http://bosswatch.labor.au>, accessed 13 August 2003.
13 Beder, pp 172–3.
14 See <http://www.afl-cio.org/aboutaflcio/wip/wip_results.cfm>
15 Hague, pp 5–6.
16 Ibid.
17 Beder, p 197.
18 Ibid, p 194.
19 M. Albert, *Parecon*, Pluto Press, London, 2003, p 90.
20 Ibid, p 90.
21 Ibid.

CHAPTER 17

1 D. Dewees, D. Duff and M. Trebilock, *Exploring the Domain of Accident Law*, 1996, pp 416–17.
2 B. Bennett, *Law and Medicine*, Lawbook Co, Sydney, 1997, p 72. In Australia, the

QAHCS study of 1995 found 16.6% of hospital admissions associated with unintended injuries or complications 'caused by health care management', 51.2% of these adverse effects were said to be 'highly preventable' and 69.6% caused death. Ibid, p 70.

3 L. Skene, *Law and Medical Practice*, Butterworths, Sydney, 2004, p 6.
4 Dewees et al, pp 416–17.
5 R.L. Abel, 'A Critique of Torts', in H. Lunz and D. Hambly, *Torts, Cases and Commentary*, 4th edition, Butterworths, Sydney 1995, p 102.
6 Lunz and Hambly, p 1.
7 Abel, p 100.
8 Fleming, p 296.
9 As Marj Milburn points out, 'bioethicists argue that … [*Bolam*, as a purely] practitioner-based standard, is ethically unacceptable … This standard completely denies the client's right to self-determination'. M. Milburn, *Informed Choice of Medical Services; Is The Law Just*?, Ashgate, Aldershot, 2001, p 29. Here indeed, doctors can 'get away with anything' merely by finding a colleague to say that they would have done the same thing in the circumstances.
10 Skene, pp 5–6.
11 Ibid, p 8.
12 Ibid, p 6.
13 Ibid.
14 K. Booth and J. Varghese, 'The Rush to Law Reform in Personal Injuries' (2003) 28 *Alternative Law Journal* 210.
15 Skene, p 13.
16 Booth and Varghese, p 211.
17 Skene, p 15.
18 Booth and Varghese, p 211.
19 Abel, p 106.
20 See S. Mann, 'Science, Corporations and the Law' (2001) 26 *Alternative Law Journal* 289.
21 Abel, p 105.
22 Dewees et al, pp 421–2.
23 Ibid, p 434.
24 Abel, p 103.

CHAPTER 18

* This chapter is written by Scott Mann and Mouaid Al-Qudah.
1 H. Hart, *Punishment and Responsibility*, Oxford University Press, Oxford, 1968, p 181.
2 Ibid.
3 Norman, On *Humanism*, pp 59–60.
4 See J.R. Searle, *Intentionality*, Cambridge University Press, 1983, p 84.
5 A. Norrie, *Crime, Reason and History*, Butterworths, London, 2001, p 121.
6 Ibid.
7 *Airedale National Health Services Trust v Bland* [1993] AC 789.
8 T.A. Mappes and D. DeGrazia, *Biomedical Ethics*, McGraw-Hill, New York, 1996, p 25.
9 Ibid, p 27.
10 Or a failure to perform a particular intentional action where there is a recognised duty to act.
11 Courts in England and Canada have associated 'sane automatism' with 'external causes' and 'insane' with 'internal causes'. But this distinction was rejected in the Australian case of *R v Falconer* (1990). Insane automatism can form the basis for a defence of insanity.

12 This is providing courts do not see the 'physically threatening situation' as a result of the action of the individual concerned. And duress is not available as a defence for the crime of murder.
13 C.M.V. Clarkson, *Understanding Criminal Law*, 2nd edition, Fontana Press, London, 1995, p 86.
14 In New South Wales, the defence of diminished responsibility was reformulated as 'substantial impairment' in 1997, with the latter understood as 'an abnormality of mind arising from an underlying condition', which was such as to 'substantially impair a person's 'capacity to understand events, or to judge whether their actions were right or wrong or to control himself or herself' at the times of acts or omissions causing death. This is supposed to exclude 'transitory disturbances of mind brought about by heightened emotions' and 'remove the requirement for a particular diagnosis of the accused's condition'. See D. Brown et al, *Criminal Laws*, p 660.
15 In the new Part 11A of the NSW *Crimes Act*, intoxication cannot be taken into account where crimes are not of specific intent.
16 Clarkson, pp 88–9.
17 Ibid, p 89.
18 Norrie, p 36.
19 Ibid, p 37.
20 Ibid, pp 38–9.
21 I. Grigg-Spall and P. Ireland, *The Critical Lawyers Handbook*, Pluto Press, 1992, p 79.
22 G. Mousourakis, *Criminal Responsibility and Partial Excuses*, Ashgate, Dartmouth, 1998, p 183.
23 Entry into the upper levels depends upon the sponsorship of those already there. And they are generally unwilling to sponsor anyone further than one level below them.
24 H. Croall, *Understanding White Collar Crime*, Open University Press, Buckingham, 2001, p 85.
25 R. Hogg and D. Brown, *Rethinking Law and Order*, Pluto Press, Sydney, 1998, pp 53–8.
26 In psychoanalytic terms, the poor and the powerless employ the defence of 'displacement' of their anger, away from those who really oppress them, who remain outside the scope of their effective action, and onto the closer targets offered by hostile peers and family members. Otherwise, the anger is turned against themselves.
27 R. Wilkinson, *Mind the Gap*, Wiedenfeld and Nicholson, London, 2000, p 24.
28 R. Hogg, 'Prisoners and the Penal Estate in Australia', in D. Brown and M. Wilkie (eds), *Prisoners As Citizens*, Federation Press, Sydney, 2002, p 14.
29 Suggesting a serious deterioration in the educational levels of prisoners since 1973. Ibid, p 15.
30 Hogg and Brown, p 69.
31 R. Hogg in Brown and Wilkie. pp 15–16.
32 Ibid, p 16.
33 A. Storr, *The Art of Psychotherapy*, Butterworth Heinemann, Oxford, 1990, p 100.
34 Ibid, p 85.
35 N. McWilliams, *Psychoanalytic Diagnosis*, Guilford, New York, 1994, p 302.
36 Ibid, pp 151–2.
37 Hayes' research in 1992 and 1995 suggested that 23.6% before the local courts in New South Wales had an IQ of less than 70, 'which placed them in … the mildly intellectually disabled category', with 14.1% functioning 'within the borderline range'. Sjogren, 'Experiences of Inmates with an Intellectual Disability', in Brown and Wilkie, p 51.
38 Ibid, p 247.
39 The 1996 Inmate Health Survey by the NSW Corrections Health Service found one-third of male and two-thirds of female prisoners tested positive for Hepatitis C. One-third of males and one-quarter of females had been confined in a juvenile institution at some time. One-fifth of men and one-quarter of the women had used heroin in

gaol. See R. Hogg, 'Prisoners and the Penal Estate in Australia', in Brown and Wilkie, pp 14–15.
40 Ibid, p 11.
41 S. Box, *Crime, Power and Mystification*, Tavistock, London, 1983, p 39.
42 Croall, p 89.
43 Chisholm and Nettheim, p 126.

CHAPTER 19

1 R. White and F. Haines, *Crime and Criminology*, Oxford University Press, Melbourne, 1996, p 142.
2 Norrie, p 25.
3 R. White and S. Perrone, *Crime and Social Control*, Oxford University Press, Melbourne, 1997, p 139.
4 Norrie, p 18.
5 Ibid, p 20.
6 Ibid, p 202.
7 Quoted in ibid, p 10.
8 Ibid, p 203.
9 Hogg in Brown and Wilkie, p 11.
10 Norrie, p 204.
11 Ibid, p 208.
12 White and Perrone, p 167.
13 Ibid, p 169.
14 Ibid.
15 Brown and Wilkie, p 18.
16 Ibid, p 6.
17 Ibid, p 197.
18 Norrie, p 214.
19 Ibid, p 125.
20 Ibid, p 215.
21 Ibid.
22 Ibid, p 216.
23 White and Perrone, p 140.
24 Ibid.
25 Ibid, p 171.

CHAPTER 20

1 For a comparison of the US and British legislation, see N. Hancock, *Terrorism and the Law in Australia: Supporting Materials*, Parliamentary Library, Research Paper No 13, 2001–02, pp 2–8.
2 'Amnesty International's Concerns Regarding Post September 11 Detentions in the USA', Amnesty International, March 2002.
3 For references and further detail see M. Head, 'The Military Call-Out Legislation – Some Legal and Constitutional Questions' (2001) 29 *Federal Law Review* 1.
4 M. Head, 'Olympic Security: Police and Military Plans for the Sydney Olympics – A Cause for Concern' (2000) 25 *Alternative Law Journal* 131.
5 Senate Foreign Affairs, Defence and Trade Legislation Committee, 'Defence Legislation Amendment (Aid to Civilian Authorities) Bill 2000', 2000, Parliament of Australia, <http://www.aph.gov.au/senate/committee/submissions/fadt_civbill.htm>; and Senate Standing Committee for the Scrutiny of Bills, *Alert Digest No 10 of 2000*, 16 August 2000.
6 It has since been revealed, however, that elite SAS personnel were deployed undercover in plain clothes, assisting the NSW police to monitor crowds during the Olympics, without approval by the Defence minister or Federal Cabinet. Cabinet's

National Security Committee subsequently approved the deployment, without any reference to the Act. See *Sydney Morning Herald*, 9 February 2001, p 6.
7 Section 51.
8 C. Doogan, 'Defence Powers Under the Constitution: Use of Troops in Aid of State Police Forces – Suppression of Terrorist Activities' (1981) 31 *Defence Force Journal* 31.
9 Ibid.
10 T. Molomby, *Spies, Bombs and the Path of Bliss*, Sydney, Potoroo Press, 1986, and J. Hocking, *Beyond Terrorism: The Development of the Australian Security State*, Sydney, Allen & Unwin, 1993.
11 W.S. Holdsworth, 'Martial Law Historically Considered' (1902) 18 *Law Quarterly Review* 117.
12 S. Greer, 'Military Intervention in Civil Disturbances: The Legal Basis Reconsidered' (1983) *Public Law* 573.
13 Quoted in B. McKinlay, *A Documentary History of the Australian Labor Movement, 1850–1975*, Drummond Publishing, Melbourne, 1979, p 377.
14 A. Blackshield, 'The Siege of Bowral – The Legal Issues' (1978) 4 *Pacific Reporter* 6.
15 Parliament of Australia, *Parliamentary Debates*, House of Representatives, 7 September 2000, p 18447.
16 Head, 'The Military Call-out Legislation'.
17 For references and further details see M. Head, '"Counter-terrorism" Laws: A Threat to Political Freedom, Civil Liberties and Constitutional Rights' (2002) 26 *Melbourne University Law Review* 266.
18 (1951) 83 CLR 1 at 187.
19 Sections 34A–34Y of the *Australian Security Intelligence Organisation Act 1979* (Cth).
20 For references and further details see M. Head, 'Another Threat to Democratic Rights: ASIO Detentions Cloaked in Secrecy' (2004) 29 *Alternative Law Journal* 127.
21 Ibid at 129.
22 Ibid.
23 Ibid.
24 This record has been documented in several works and official inquiries. See for example D. McKnight, *Australia's Spies and their Secrets*, Allen & Unwin, Sydney, 1994; R. Hall, *The Secret State*, Cassell Australia, Sydney, 1978; F. Cain, *The Origins of Political Surveillance in Australia*, Angus & Robertson, Sydney, 1983; F. Cain, *ASIO, an Unofficial History*, Melbourne, Spectrum, 1994; J. Hocking, *Beyond Terrorism: The Development of the Australian Security State*, Allen & Unwin, Sydney 1993; Commonwealth of Australia, *Royal Commission on Intelligence and Security: Fourth Report*, Vols 1 and 2, AGPS, Canberra, 1977.

CHAPTER 21

1 J. Morrison, 'The Trafficking and Smuggling of Refugees: The End Game in European Asylum Policy?, UNHCR, Geneva, July 2000, <http://www.unhcr.ch/evaluate/reports/traffick.pdf>.
2 M. Head, 'The High Court and the Tampa Refugees: A Critical Examination of *Vadarlis v Minister for Immigration and Multicultural Affairs*' (2002) 11 *Griffith Law Review* 23.
3 The statistics in this section are sourced from M. Head, 'Refugees, Global Inequality and a New Concept of Global Citizenship' [2002] *Australian International Law Journal* 57 at 58–9.
4 J. Harding, *The Uninvited: Refugees at the Rich Man's Gate*, Profile Books, London, 2000.
5 E. Feller, 'The Convention at 50: The Way Ahead for Refugee Protection' (2001) 10

Forced Migration Review 6 at 7.

6 Department of Immigration and Multicultural Affairs, *Fact Sheet 1: Immigration – The Background*, 23 June 2000, p 1.

7 Commonwealth Parliamentary Debates, House of Representatives, Vol. 1, p 5153, 25 September 1901, extracted in B. McKinlay, *A Documentary History of the Australian Labor Movement, 1850–1975,* Drummond Publishing, Melbourne, 1979, p 28.

8 D. North, *Capital, Labor and the Nation-State*, Labor Publications, Detroit, 1992, p 1.

9 J. Hathaway, *The Law of Refugee Status*, Toronto, Butterworths, 1991, pp 10–11.

10 Per Gummow J in *Minister for Immigration and Multicultural Affairs v Ibrahim* (2000) 175 ALR 585 at [137].

11 M. Crock, *Immigration and Refugee Law in Australia*, Federation Press, Sydney, 1998, p 32.

12 J. Hathaway, 'Can International Refugee Law be Made Relevant Again?', in J. Hathaway (ed), *Reconceiving International Refugee Law*, Martinus Nijhoff Publishers, The Hague, 1997, p xxix.

13 G. Van Kessel, 'Global Migration and Asylum' (2001) 10 *Forced Migration Review* 10.

14 A. Dummett and A. Nicol, *Subjects, Citizens, Aliens and Others, Nationality and Immigration Law*, Weidenfeld and Nicholson, London, 1990, p 13.

15 *Lim v Minister for Immigration, Local Government and Ethnic Affairs* (1992) 176 CLR 1.

16 Dummett and Nicol, pp 11-12.

17 Ibid, p 13.

18 R. Falk, 'The Making of Global Citizenship' in B. Van Steenbergen (ed), *The Condition of Citizenship*, Sage, London, 1994, p 139.

19 A. Negri and M Hardt, *Empire*, Harvard University Press, Cambridge, 2000.

INDEX

LAW IN PERSPECTIVE is intended to encourage critical, responsible and creative thinking about law as a system of ideas and a social institution. Written partly as a textbook for first-year law students, it explores the relationships between law, logic and science; examines the socio-economic role of law; and asks what role the legal system plays in alleviating or exacerbating contemporary social problems (such as crime and punishment, terrorism, refugees and tort law reform).

To place law in its historical, philosophical, economic, political and social contexts, Michael Head and Scott Mann focus upon a range of powerful critical thinking tools in the form of ideas and techniques drawn from logic, science, ethics, and political and social theory.

LAW IN PERSPECTIVE provides an accessible, concise and practical introduction to the fundamental ideas of

- good and bad reasoning
- scientific method
- ethical theories and concepts
- politics and political economy.

UNSW PRESS
ISBN 0-86840-834-4
9 780868 408347