

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

A Thesis Submitted for the Degree of PhD at the University of Warwick

<http://go.warwick.ac.uk/wrap/60433>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

**Tort From Scratch: The
Philosophical Foundations of Harm,
Actionability and Corrective Duties**

Adam Slavny

Contents

Chapter 1. Introduction	8
1. Unconditional Instrumentalism in Tort Theory	12
A. Economic Theories	14
2. Conditionally Instrumentalist and Non-Instrumentalist Theories	19
A. Richard Epstein's Theory of Strict Liability	19
B. Rights-based Theories	24
C. Civil Recourse Theory	29
D. Concluding Remarks	32
3. Normative Tort Theory and Tort Law	34
Chapter 2. Methodology and Outline	40
1. Corrective Justice	40
A. Jules Coleman	41
B. Ernest Weinrib	43
C. The Defence of Corrective Justice	46
D. John Gardner	50
2. Development of the Thesis	52
3. Normative Methodology	55
Chapter 3. Harm	65
1. The Currency of Harm	68
2. Defining Setbacks	77
A. Criticisms of the Counterfactual View	77
B. Event-based Accounts of Harm	80
C. The Complex Counterfactual View	88
3. Conclusion	97

Chapter 4. Corrective Action	100
1. Previous Accounts of Corrective Measures	104
2. Negating and Counterbalancing	106
3. Primary and Secondary Duties	111
A. Preventive and Non-Preventive, Forward and Backward Looking	111
B. The Justifiability of Secondary Duties	113
4. What Corrective Duties May Exist?	116
5. Identically Off and As Well Off	119
6. Negating, Counterbalancing and Remedies in Private Law	125
7. Choice and the Normative Priority of Negating over Counterbalancing	128
8. Conclusion	129
Chapter 5. Actionability	132
1. Compensation and Dworkinian Liberalism	141
2. Defending <i>The Respect View</i> Against <i>The Objective View</i>	144
3. Quantifying Compensation	150
4. Defending <i>The Objective View</i> against <i>The Respect View</i>	153
5. The Bad Faith Objection	160
6. Conclusion: Arriving at <i>The Mixed View</i>	164
Chapter 6. Corrective Duties	166
1. The Three Principles	171
A. Nonreciprocity	171
B. Avoidability	175
C. Benefit	180
2. The Moral Underpinning of Nonreciprocity	181
3. Nonreciprocity and Avoidability	186

4. Nonreciprocity and Benefit	189
5. The Components of Nonreciprocity	193
6. Conclusion	199
Chapter 7. Corrective and Distributive Justice	202
1. What is Corrective Justice?	202
2. Corrective and Non-Corrective, Agent-Centred and Agent-Neutral	208
3. The Relationship Between Distributive and Corrective Justice	214
4. Does Corrective Justice Necessarily Raise Issues of Distributive Justice?	218
5. The Relevance of Distributive Arguments to Standards of Liability	222
6. Conclusion	232
Chapter 8. Conclusion	234
Bibliography	238

Acknowledgements

This thesis has been enabled and shaped by many people. I have had opportunities to present my work at a couple of conferences organised by the Warwick law school to a receptive and thoughtful audience of colleagues. The Centre for Ethics, Law and Public Affairs was also an invaluable forum in which I presented an early version of chapter 3. The cheerfully contentious atmosphere of CELPA helped to introduce me to the world of normative philosophy. Thanks in particular to Matthew Clayton, the organiser of CELPA, who gave written feedback and has discussed my work with me on other occasions.

Thanks also to the participants of CELPA and others in the university with whom I have discussed issues of philosophy and law, especially Kimberly Brownlee and Thomas Parr who gave me written feedback on my work. I also thank Tim Fowler and Tom for our many helpful conversations. I offer Tom an apology as well: I have shamelessly benefitted from your tragic inability to talk about anything other than darts and philosophy.

Special thanks should go to Paul Raffield, who was an inspirational teacher when I was an undergraduate and a source of encouragement and support when I returned for doctoral studies. He has helped me through a number of tricky situations: diving head-first into teaching, funding a PhD, and pursuing my research interests, even if they diverge from his own. Without him I would probably not have started, let alone completed, a PhD.

This thesis owes a huge debt to Victor Tadros, from whom I have learnt a great deal about philosophy (and almost nothing about law). He is a constant source of good ideas, many of which have found their way into the present work. His skill, creativity and attitude to philosophy have influenced and benefited me greatly. I am also grateful for his humour and support throughout.

I have a great debt to my family and friends. My parents, being generally bookish and curious, gave me the initial impetus for further study; my brother and sister – Peter and Rachel – are partly responsible for my entrance into philosophy by showing me that science involves too many graphs; and many of my friends have offered stimulating discussion over

the last three years. Finally, thank you to Alice Leonard, for her kindness, levity, warmth and intelligence.

Declaration

Chapter 4 substantially overlaps with “Negating and Counterbalancing: A Fundamental Distinction in the Concept of a Corrective Duty” forthcoming in *Law and Philosophy*. This thesis is my own work and has not been submitted for a degree at another university.

Abstract

This thesis builds a normative theory of tort law by exploring the philosophical foundations of harm, actionability and corrective duties. In chapters 1 and 2 I survey previous literature in tort theory, arguing that normative questions have generally been neglected in favour of interpretive ones. I also defend the case-based methodology, familiar to moral philosophers, which I employ throughout. In chapter 3 I investigate the metaphysics of harm, making two claims: first, we should define harm as setback to wellbeing, and second, we should accept a complex version of the counterfactual view. In chapter 4 I distinguish between two fundamental forms of corrective action – negating and counterbalancing – and argue that they have important implications for tort theory. In chapter 5 I inquire whether a victim’s false beliefs about her wellbeing should have any impact on her claim to compensation against a wrongdoer. Chapter 6 offers a critique of George Fletcher’s theory of reciprocity as a moral basis for corrective duties. Having rejected it, I propose a set of alternative principles that more plausibly explains our judgement about whether an injurer ought to compensate her victim. Finally, chapter 7 discusses the relationship between corrective and distributive justice. I argue that, contrary to the claims of some theorists, corrective justice cannot be insulated from distributive justice.

Introduction

This is a study in normative tort theory. The first difficulty facing such a study is that the field of normative tort theory does not exist. That is to say, it does not exist as a distinct area of philosophical scholarship, as a systematic and rigorous study of the normative issues raised by tort law. There is, nevertheless, a rich variety of theoretical outlooks in tort. Most notable is the debate between the economic theorists and proponents of corrective justice, but other approaches such as those grounded in the social contract tradition, libertarian values, and more recently rights and civil recourse, also populate the field. But these contributions are primarily interpretive rather than normative: they seek to rationally reconstruct the concepts, decisions and patterns of reasoning embodied in tort practice, rather than arguing how tort law ought to be arranged, or whether it ought to exist at all.¹

In this respect, the current state of tort theory differs from the philosophy of criminal law. To be sure, interpretive questions are debated in both disciplines. It is a staple of theoretical legal scholarship to construct explanatory rationales for various areas of law. But in the philosophy of criminal law, normative issues are at the top of the agenda. Much ink has been spilt over the basis on which conduct can be legitimately criminalised and the justification of state punishment, to take two central examples. Justifications of punishment or principles delimiting the proper scope of criminalisation, if they can be found, do not wholly depend on any particular explanation of what the criminal justice system is doing in practice.²

¹ For a well-known discussion of descriptive, historical, interpretive and normative theories of law, see Stephen Smith, *Contract Theory*, Oxford: Clarendon Press, (2004), pp. 3 – 6.

² Although the point is only illustrative, it is worth noting that this is controversial. It might be argued that a normative theory must take into account the best explanation of the criminal law, or that there is no right answer to the normative question independently of the historical, sociological or

One can only speculate as to why this difference has emerged. Perhaps it is because the concepts that figure in criminal law, such as justification, excuse and intentional harm, raise moral problems more explicitly. But as we will see throughout this thesis, tort law also touches upon foundational issues in moral and political philosophy. Perhaps it is because scholars have been content with partially normative theories – those that treat justification as a secondary or complementary aim to the main business of interpretation. If one assumes that tort law as it currently exists is justified, we need do no more than extract its moral principles: interpretation will *constitute* justification. Obviously enough, this assumption is question-begging. We cannot be satisfied that modern tort law is morally acceptable in advance of a full justification of it.

This paucity of analysis has placed us in a startling predicament: we have *no idea* whether tort law is justified, which doctrines are defensible and which should be reformulated or discarded, what principles explain these conclusions, and so on. The purpose of this thesis is to take a small step beyond this disconcerting state of affairs.

In broad outline, here is how the thesis will progress (I leave a more detailed summary until the next chapter, where we will also deal with the methodology). We will be mainly concerned with the conditions under which one person owes a corrective duty to another, typically in cases where the first person causes harm to the second. Before we investigate these conditions, a number of prerequisites must be addressed. First, if corrective duties are usually owed in respect of *harm*, we need to know what harm is and how it can be quantified. Resolving these questions will deepen our understanding of harm as a moral concept, explaining its importance for corrective duties (which will presumably emerge from its importance generally), and tell us how to make judgements about harm in difficult cases such as those involving pre-emption. Second, we must not be content with a vague reference

other explanatory context within which criminal law practice must be understood. For a brief critical discussion of these ideas in relation to criminal responsibility, see Victor Tadros, *Criminal Responsibility*, Oxford: Oxford University Press, (2005), pp. 3 – 8.

to 'secondary', 'remedial' or 'corrective' duties, and explore different types of corrective action. I argue that there is a distinction between negating and counterbalancing actions, which is significant in that it has implications for permissibility and other moral judgements, and can be obscured by the use of generic terms.

Once this is completed, I investigate the moral foundation of corrective duties. I undertake this through a critical analysis of a promising set of theories that appeal to reciprocity between potential injurers and victims, originally proposed by George Fletcher and amended by Gregory Keating. Although I reject these views, examining them allows us to reveal a more plausible set of principles, which, I tentatively argue, constitute the moral basis of corrective duties. These principles are rooted in the avoidability of harm, the benefits generated by harmful activities, and considerations of distributive fairness.

Even if it is established that a person caused harm by committing the kind of conduct that renders her liable, it is a further question whether a corrective duty ought to be imposed on her, and if so what the content of that duty ought to be. In other words, it is a further question whether the consequences of that conduct are *actionable*. Actionability has been given very little attention by tort theorists, but it raises interesting problems. The problem that I address is how we should respond in cases where a victim has false beliefs about her own wellbeing, and therefore forms false beliefs about whether she is harmed by wrongful conduct. If she is harmed but falsely believes she is benefitted, should her injurer offer compensation? What if she is benefitted but falsely believes she is harmed? I defend a view that strikes a balance between two desirable goals: respecting a victim's autonomous capacity to form her own beliefs about wellbeing and compensating harm for which an agent is responsible.

In the final section I explore the relationship between corrective and distributive justice. I challenge the arguments of some prominent tort theorists that the two are independent of one another. Mounting this challenge is also necessary to lend support to my own arguments in chapter 6 that distributive considerations are relevant to the moral basis of

corrective duties. I explore the relationship between corrective duties and other sorts of duties, concluding that specifically corrective duties are not of special moral concern.

Before we get to this, however, the first two chapters will give some background to tort theory, outline the thesis in greater detail and defend the normative methodology employed throughout. Section 1 of the introduction will survey pre-existing theories in the philosophy of tort. The aim of this survey is not to provide a complete inventory of current theories or to conduct a comprehensive critical analysis. Instead, I will highlight areas where moral argument is incomplete or ineffective. This will contextualise the remainder of the thesis, and also lend some credence to my bold claim that normative tort theory does not exist. In section 2 I will say more about the relationship between normative tort theory and tort law. Inevitably, the former is both broader and narrower than the latter. It is reasonable to expect that the important ethical questions do not map perfectly onto current taxonomies in private law. But with that said, some of these philosophical problems, such as the moral basis of liability to compensate for harm that one agent causes another, also represent paradigmatic cases in tort law.

The survey of pre-existing literature in the present chapter omits Fletcher and Keating's reciprocity theories and the corrective justice theorists. The former are worth assessing solely for their moral appeal, so I will save them for our investigation of the basis of corrective duties in chapter 6. I discuss corrective justice theories in chapter 2. Given the sustained attention to corrective duties in this thesis, one might assume that my approach is broadly within the corrective justice tradition. But if this label is to be applied, however loosely, some fundamental differences must be noted. Two in particular are worth mentioning at the outset. First, the arguments offered here are purely normative, while most corrective justice theorists adopt a mixed methodology, as I will explain in chapter 2. Second, some corrective justice theorists do not defend any substantive principles that explain what duties we have, and those that do offer different principles to those I will advocate later. With these qualifications made, the arguments I will develop can be seen as giving shape to a normative theory of corrective justice.

1. Unconditional Instrumentalism in Tort Theory

Unconditional instrumentalist theories, as I will understand them here, hold that tort law is explicable and/or justified only in virtue of its efficacy in achieving some independent aim. The duties of tort law may not refer to these instrumental goals, but they are fully justified or explained by them. These forms of instrumentalism are unconditional because they do not place any constraints on the pursuit of these goals: tort law is determined entirely by how best to achieve them. The relevant aims can be monistic or pluralistic. We will consider a monistic form of unconditional instrumentalism, whose aim is the maximisation of wealth (or more accurately the minimisation of accident costs), shortly. If that theory specified a further goal – deterring wrongdoing, for instance – it would be pluralistic in its aims. What unites these theories is their claim that we should not acknowledge any moral constraints on tort practice which limit its pursuit of instrumentalist aims.

Conditional instrumentalist theories hold that tort law pursues independent goals, or is permitted to pursue them, but only subject to certain constraints.³ Sometimes ancillary benefits are produced as a by-product of enforcing duties that are justified on non-instrumentalist grounds. One example of this might be that enforcing corrective duties deters people from committing future wrongs. This, in itself, does not make a theory conditionally instrumentalist, since the primary aim is the enforcement of the duty for its own sake, rather than deterrence. If the instrumental benefit is an irrelevant by-product as far as the theory is concerned, it can hardly be instrumentalist in any meaningful sense. But instrumental goals

³ For an explanation of conditional and unconditional forms of instrumentalism, as well as the relationship between instrumentalist theories and consequentialist or non-consequentialist moral frameworks, see Victor Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law*, Oxford: Oxford University Press, (2011), especially chapter 2.

may play a more prominent role in the theory. For example, they may operate as a decision mechanism when there are insufficient resources to enforce all equivalent corrective duties. If there are only resources to enforce one of two corrective duties that are equivalent in every respect except that one will have greater deterrent effects, we might think that it is permissible to enforce the duty that will best further the goal of deterrence. This decision is made for an instrumental purpose, but it is only permissible to use those who already owe enforceable duties to pursue this goal. It would be wrong, on these theories, to use someone who is not a duty-bearer for this end.

Finally, *non-instrumentalist theories* hold that tort law cannot be explained or justified with reference to any of its instrumental effects. Instrumental goals cannot even play a secondary role as a decision mechanism. Normative non-instrumentalist theories insist that if the law is allowed to enforce corrective duties, this permission is justified purely on non-instrumentalist grounds. Explanatory versions will have a hard time accommodating the fact that policy plays an explicit role in both statutory law and judicial decision making.⁴ But it might be claimed that the rare examples of instrumentalism in the law are anathema to its general unifying principles, which are resolutely non-instrumentalist.

It is not always straightforward to tell which label best applies to a given theory. Economic and utilitarian theories are unconditionally instrumentalist as they evaluate the law only in terms of its efficacy in maximising wealth or utility. At the other end of the spectrum, Ernest Weinrib's interpretation of corrective justice is emphatically non-instrumentalist, as I will explain in the next chapter. Others are harder to categorise, often because they fail to comment on the role of deterrence or policy within the theory. We will begin our survey with the economic theory – the most well-known form of unconditional

⁴ For an example of instrumentalism in the statute books, see the Compensation Act 2006. Section 1 states that when deciding whether a party has fallen below the relevant standard of care for negligence or breach of statutory duty, a court can have regard to whether the relevant precautions might prevent desirable activities being undertaken.

instrumentalism. We will then proceed to other theories centring on causation, rights and civil recourse. All of these are hostile towards instrumentalism, but they usually fail to specify whether they are conditionally instrumentalist or non-instrumentalist, and some can be formulated into both molds.

A. Economic Theories

Economic theories are unconditionally instrumentalist because they view tort law solely as a mechanism for reducing the cost of accidents and their prevention. Before outlining some central tenets of economic theories, let us introduce three concepts necessary to understand them: Pareto improvement, Pareto optimality and Kaldor-Hicks efficiency. The concept of Pareto improvement supplies a way of ranking states of affairs. One state of affairs, A, is a Pareto improvement over another, B, if the move from B to A makes no one worse off and at least one person better off. The criterion assumes that a person is better off in A than B if she prefers A to B and worse off if she prefers B to A. It also assumes that she is made neither worse off nor better off by being in A or B if she is indifferent between them.

Repeated Pareto improvements lead to Pareto optimality, defined as that state of affairs in which no further Pareto improvements can be made. Any shift to a different state of affairs will make at least one person worse off. In practice, however, it is almost impossible to take some social action without making at least one person worse off. Kaldor-Hicks efficiency, by contrast, permits changes in which individuals are made worse off, as long as those who are made better off could compensate the losers in principle.⁵ It is not necessary that the winners actually compensate the losers, only that they have profited

⁵ See Nicholas Kaldor, "Welfare Propositions of Economics and Interpersonal Comparisons of Utility", *The Economic Journal*, 49, (1939), 549 – 552; John Hicks, "The Foundations of Welfare Economics", *The Economic Journal*, 49, (1939), 696 – 712.

enough from the transaction that they *could* do so. Kaldor-Hicks efficiency is thus applicable in a much broader range of circumstances, but since it does not demand actual compensation, it usually makes at least one party worse off.

So how are these concepts employed in economic theories of law? One foundational text in the economic movement is Ronald Coase's *The Problem of Social Cost*.⁶ Put briefly, Coase showed that assuming rational behaviour, no wealth effects and no transaction costs, the prior allocation of legal entitlements will not affect the efficient distribution of resources. Regardless of who is initially assigned the right to use a patch of land, for example, it will eventually reside with its higher valuer. As there are no transaction costs, the highest valuer will purchase the entitlement from its owner, or if the person to whom it is assigned is the highest valuer, she will refuse to sell.

This outcome is Pareto optimal because there is no further transaction that will make at least one person better off and no one worse off. The land right now resides with its highest valuer and any further transfer will make her worse off. Coase's result assumes that there are no transaction costs – an umbrella term that encompasses identifying parties with whom to transact, negotiating, drawing up contracts, and so on. Other key proponents of the law and economics movement, such as Richard Posner, are interested in the legal assignment of rights under conditions of *high* transaction cost. Several common scenarios in tort law, such as accidents on the road and in the workplace, are characterised by high transaction costs.

Posner argues that in these contexts the law should intervene to obtain the result that would be produced if there were no transaction costs, thus mimicking the market. This circumvents the restrictions imposed by the Pareto criterion by relying on the notion of Kaldor-Hicks efficiency. In Coase's model, each move is a Pareto improvement since neither party to the transaction is rendered worse off, and at least one party secures a gain. If B is not compensated by A for transferring her entitlement, B can simply veto the transaction. Kaldor-Hicks efficiency only requires that A gains more than B loses. On

⁶ Ronald Coase, "The Problem of Social Cost", *Journal of Law and Economics*, 3(1), (1960), 1 – 44.

Posner's model, the law assigns the entitlement *directly* to A, avoiding the need for A to purchase it from B. This maximises wealth, but it leaves B worse off than would the Pareto criterion.

In *The Economic Structure of Tort Law*, Posner and William Landes argue that the common law of torts is consistent with economic principles of wealth maximisation, understood in terms of Kaldor-Hicks efficiency. They argue that the Learned Hand formula for the standard of care, in which a person falls below the designated standard if she fails to take cost-justified precautions, conforms to the economic model, as do defences, strict liability for ultra-hazardous activities, vicarious liability, intentional torts and the liability of joint and multiple tortfeasors. These are all explanatory claims, and they are well summed up by the opening sentence of the book: "the common law of torts is best explained as if the judges who created the law through decisions operating as precedents in later cases were trying to promote efficient resource allocation".⁷ This project stands or falls depending on how well the hypothesis that judges seek to promote economic efficiency explains the evolution and current state of tort law.

In other work, Posner claims that this explanation even extends to moral obligations and the attitudes of disapproval provoked by their breach. In *The Economics of Justice*, Posner argues that economic efficiency defines the word 'justice' itself, at least as applied to the law.⁸ He even contends that economic efficiency can explain the legal and other social institutions of primitive or pre-literate societies, such as the governments depicted by Homer in *The Iliad* and *The Odyssey*.⁹ These claims are extravagant and have been criticised and

⁷ See William Landes and Richard Posner, *The Economic Structure of Tort Law*, Cambridge, MASS: Harvard University Press, (1987), p. 1.

⁸ See Richard Posner, *The Economics of Justice*, Cambridge, MASS: Harvard University Press, (1981) and "A Theory of Negligence", *Journal of Legal Studies*, 1, (1972), 29 – 96, p. 33.

⁹ Richard Posner, *The Economics of Justice*, pp. 119 – 145.

rejected by others.¹⁰ But they still do not amount to a normative defence of economic efficiency. Even if economic efficiency captures a common meaning of the word ‘justice’, this does not show that the employment of this principle to guide legal decision-making is justified. To assume otherwise would be substantially to assimilate justification to explanation. It cannot be true that to explain the concept of justice that predominates in our legal system, and has fashioned legal systems throughout history, is also to justify it, any more than explaining the theological rationale for traditional religious practices justifies their continuation.

In general, relatively little has been offered in the way of justification for the presuppositions of the economic theory. Steven Shavell, one of its foremost proponents, endorses the maximisation of social welfare as a guiding moral principle, but elides a sustained defence of it. He concedes that other ‘deontic’ moral concepts ought to be incorporated into the economic theory, but only because and to the extent that they tend to advance overall welfare. They do this by curbing socially undesirable selfishness and prevent us acting in ways that are tempting in the short term but contrary to our interests in the long run.¹¹ This shows that one who is already committed to maximising social welfare has reason to absorb other moral notions that are instrumental in achieving this aim, but this is not a justification of social welfare against competing views. Shavell’s conclusion to this strikingly incomplete discussion has a dismissive air. He states that welfare economics “seems to provide an intellectually attractive and generally satisfactory lens for understanding and analyzing morality and law... Although I realize that this view is in

¹⁰ See Ronald Dworkin, *A Matter of Principle*, London: Duckworth, (1985), Ch. 12 and *Law’s Empire*, London: Fontana, (1986), Ch. 8.

¹¹ See Steven Shavell, *Foundations of Economic Analysis of Law*, Cambridge, MASS: Harvard University Press, (2004), pp. 595 – 612 and Louis Kaplow and Steven Shavell, *Fairness Versus Welfare*, Cambridge, MASS: Harvard University Press, (2002).

tension with the great weight of legal writing and thinking, it is the only one that I can comfortably endorse.”¹²

Perhaps the most extensive normative defence of the economic theory is the argument for wealth maximisation offered by Posner in *The Economics of Justice*.¹³ This defence consists largely of critiques of other views, including utilitarianism and those developed by John Rawls and Ronald Dworkin. Posner’s analysis is sweeping and a number of his points are far too quick to be convincing. For example, he notes that Rawls and others have promoted the view that an individual’s genetic endowment is a kind of accident, lacking in moral significance. He asserts that this is inconsistent with the Kantian conception of individuality from which the view purports to derive, for two reasons. First, it does not take seriously the differences between persons, and second, a policy of redistribution impairs the autonomy of those from whom the redistribution is made.¹⁴ The first point is really a restatement of his claim rather than a defence of it, and the second is inconclusive. Before we can agree that redistributing wealth from the talented is an unacceptable violation of autonomy, we need to know much more about the value of autonomy, how it ought to be respected, and how things that enhance or restrict autonomy, such as resources, opportunities, right and duties, ought to be distributed. For all he argues to the contrary,

¹² Shavell, *Foundations of Economic Analysis of Law*, at p. 646.

¹³ Posner’s arguments have been criticised quite extensively, and I will not attempt to rehearse the criticisms here. See Jules Coleman, “The Normative Basis of Economic Analysis: A Critical Review of Richard Posner’s ‘The Economics of Justice’”, *Stanford Law Review*, 34(5), (1982), 1105 – 1131; Ronald Dworkin, *A Matter of Principle*, London: Duckworth, (1985), Ch. 12 and *Law’s Empire*, London: Fontana, (1986), Ch. 8.

¹⁴ Posner, *The Economics of Justice*, p. 76.

Posner's comments might also misrepresent Kantian philosophy (in any case, he elsewhere dismisses the value of Kantian ethics).¹⁵

I will not offer any further critique of Posner's defence of wealth maximisation here. Instead, my purpose is to emphasise the general lack of attention paid to the moral defence of economic efficiency as the only legitimate purpose of tort law. The fact that Posner's defence, incomplete though it is, stands out as a rare attempt to justify the economic theory, only serves to highlight this lacuna.

2. Conditionally Instrumentalist and Non-Instrumentalist Theories

A. Richard Epstein's Theory of Strict Liability¹⁶

There are a group of theories that ground liability in one agent's causing harm to another. These theories are not unconditionally instrumentalist as they take this condition to be sufficient to justify a corrective duty. If causation is the main ground of liability, instrumentalist factors are not of prime importance. Whatever they may say about the

¹⁵ See Posner, *The Problems of Jurisprudence*, Cambridge, MASS: Harvard University Press, (1990), p. 329. Posner claims that "the ethics of Kant are too abstract, however, to guide the design of legal doctrines."

¹⁶ Although I only have space to address Epstein's theory, others have also held that causation is central to liability in tort. See Richard Wright, "Causation in Tort Law", *California Law Review*, 73, (1985), 1735 – 1828; H. L. A. Hart and Tony Honoré, *Causation in the Law*, 2nd Ed., Oxford: Clarendon Press, (1985); A. M. Honoré, "Responsibility and Luck: The Moral Basis of Strict Liability", *Law Quarterly Review*, 104, (1988), 530 – 52; and A. M. Honoré, "Ownership", in *Oxford Essays in Jurisprudence*, Ed., Anthony Gordon Guest, Oxford: Oxford University Press, (1961).

conditional pursuit of instrumentalist goals, these goals are not the central justification or explanation for tort law. One such view is Richard Epstein's theory of strict liability.

Unlike Posner and Shavell, Richard Epstein does not attempt to explain tort law. He rejects 'moral' explanations of negligence because that they cannot show why the standard of care is insensitive to the individual's capacity to conform to it.¹⁷ If negligence imposes a moral standard, excuses such as insanity ought to be available, since a person cannot be held morally responsible for her actions if she is unable to appreciate their nature and consequences.¹⁸ As far as explanation is concerned, Epstein nods approvingly at the economic argument, referring to the cost-benefit analysis implicit in the Learned Hand formula for the standard of care. Instead, he rejects the principles of negligence on normative grounds.

The essence of Epstein's theory is that the proposition 'A harms B', once adequately clarified, is a suitable justification for the imposition of liability. The standard method for establishing that A caused harm to B is the application of a simple counterfactual test. The court asks the following question: 'Would the claimant have suffered harm but for the tort committed by the defendant?' If the answer is no, the defendant is shown to have caused the claimant's loss. Epstein is aware that the 'but for' test is over-inclusive. If A unfairly dismisses her employee, B, who drives home and negligently injures C, the unfair dismissal by A is, according to the 'but for' test, a cause of C's injury. But the intuitive and legally accurate conclusion is that B's negligence rather than A's wrong is the cause of C's injury.

¹⁷ Richard Epstein, "A Theory of Strict Liability", *The Journal of Legal Studies*, 2(1), (1973), 151 – 204, p. 153.

¹⁸ Epstein is relying on a particular meaning of the word 'moral' here: a moral standard is one that is fully sensitive to the blameworthiness of the injurer. But there are other conceptions of responsibility that do not depend on blameworthiness and are equally grounded in 'moral' considerations. In fact, Epstein's own theory provides one such conception.

Despite this dilemma, Epstein agrees with Hart and Honoré¹⁹ that it is a *non sequitur* to conclude that causation has no useful meaning and is just a disguise for judicial discretion or policy-making. He seeks to show that the concept of causation, when properly analysed, is internally coherent and relevant to the ultimate question of who should bear the loss”.²⁰

Epstein proceeds by offering four paradigmatic cases of causation: ‘A hit B’, ‘A frightened B’, ‘A compelled B to hit C’ and ‘A created a dangerous condition resulting in harm to B’. Any of these generic scenarios raise a *prima facie* case against A. He takes pains to refine each category so that they exclude cases that are not plausibly actionable. Consider two examples posed by Epstein. In the first, A raises his hand in an objectively non-threatening manner but B, who is of a very nervous disposition, gets frightened. In the second, A leaves her carving knife in a drawer where it is stolen by someone who uses it to harm B.²¹ These cases seem to fit the paradigms for frightening and creating dangerous conditions, but it is not plausible that A owes compensation in either case. In order to avoid these implications Epstein claims that in the first case B induces her *own* fright, and of the second he insists that a defendant must create a *dangerous* condition rather than a mere necessary factual condition.

In refining these categories, Epstein ‘moralises’ causation by stipulating a series of conditions as internal to causal requirements that are designed to exclude implausible claims. Perhaps Epstein’s judgements about liability are intuitive, but it is hard to avoid the impression that his theory of causation has become *ad hoc*. If we are interested in causation then why should we ignore A’s role in causing B’s fright, when A’s action was factually necessary for B’s harm and stemmed from his agency? We must bear in mind that Epstein’s theory purports only to be an account of causation. He cannot help himself to the conclusion that his account is preferable to others because it has more plausible implications about

¹⁹ See H. L. A. Hart and A. M. Honoré, *Causation in the Law*, p. 3.

²⁰ Epstein, “A Theory of Strict Liability”, p. 165.

²¹ See *Ibid.* pp. 172 and 179 respectively.

liability when it is *he* who has introduced the controversial assumption that mere causation is sufficient for liability.

As Epstein later recognised, his theory must presuppose some account of rights or self-ownership.²² In future writing, he placed greater emphasis on the libertarian value of self-ownership, holding that “every personal injury action... rests on an unspoken assumption that each person owns his own body.”²³ This claim is still too strong, however. Self-ownership is one possible ground for personal injury actions but it is not the only candidate. An action could equally be motivated by harm, defined as setback to wellbeing, even if no one owns their own body. A person who believes that all human bodies are owned by God could still conceivably recognise personal injuries as wrongs committed against other individuals, because damaging their bodies harms them.

In any case, what is now required is a comprehensive defence of self-ownership as the appropriate basis for the law of torts. Unfortunately, this is not provided. Epstein states that “ownership is typically defined in terms of inviolability that in turn suggests absolute protection against all invasions”.²⁴ He notes that support for this statement is found in many areas of private law. For example, specific performance is the remedy for breach of land-sale contracts even if the seller is prepared to pay full damages to keep the land. He adduces other examples which we need not detail here. The problem with the argument is evident: just when Epstein needs to give his theory a strong normative foundation, he slips into an explanatory mode. Even if it is true that certain rights in private law entail absolute protection against invasion, this does not justify the proposition that tort law ought to impose a general regime of strict liability. To do this, Epstein must do more than show that his

²² See Richard Epstein, "Nuisance Law: Corrective Justice and Its Utilitarian Constraints," *The Journal of Legal Studies*, 8(1), (1979), 49 – 102.

²³ *Ibid.* p. 50.

²⁴ Richard Epstein, "Causation and Corrective Justice: A Reply to Two Critics," *The Journal of Legal Studies* 8, (1979), 477 – 504, p. 500.

position is consistent with several decided cases. He must demonstrate the value of protecting this system of rights and show that it outweighs competing considerations. This will involve answering a number of difficult questions. Will a blanket regime of strict liability discourage desirable activities? Is it fair to hold negligent and non-negligent parties equally liable for accidental harm? Given that it is harder to avoid causing harm non-negligently than negligently, is it acceptable that liability is so sensitive to bad luck?

Moreover, the libertarian values on which the theory is based are not fully clarified. Epstein sometimes caches out self-ownership in terms of exclusive control or sovereignty, but this does not tell us exactly which bodily invasions are actionable. Consider the case in which A negligently leaves an open pit in the street and B falls into it. A ought to be liable on Epstein's view, but it is not clear how B's sovereignty has been compromised. She retains exclusive control over her own body and there are no external forces that manipulate her voluntary movements. Epstein would say that A is liable because her actions were not a mere factual condition of B's injury – they created a dangerous situation. But B's sovereignty seems no more affected in this case than it does in those where A's actions are a mere necessary factual condition. Imagine that A picks up a stone on a whim and puts it down again, a few inches from where she found it. B then trips over the stone and suffers an injury. It is hard to see how the difference between these cases has anything to do with the sovereignty of the victim.

Before moving on, let me make a qualification. Even if we reject Epstein's strong thesis about causation, it does not follow that causation is morally irrelevant: it might take a lesser role in the best overall account. Owing to constraints of space, I will not discuss this possibility any further. It is worth registering my belief that mere causation (independent of fault or reasonable foreseeability) is morally relevant and can affect the corrective duties people incur when they harm others. But the defence of this view must be left for another time.

Like Posner, Epstein attaches his arguments to a broader set of moral claims. Whereas Posner's theory relies on the appeal of wealth maximisation, Epstein's is dependant

on rights of self-ownership. But ultimately he fails to adequately refine or defend these foundations. It is surprising that even Epstein's theory – the most explicitly normative of those considered so far – falls short of delivering a satisfactory justification of the moral assumptions on which it depends.

B. Rights-based Theories

In recent years, rights-based theories have gained popularity in private law scholarship. Most are non-instrumentalist as they deny that it is ever legitimate for tort law to pursue instrumental goals. This would be to misunderstand the nature of rights, these theories claim, which supply the overarching framework for private law. Rights theories are not exclusively non-instrumentalist, however. Some are conditional as they allow policy reasoning to play a limited role in judicial decision-making.²⁵ But all rights theories share some central claims, which are summed up by Robert Stevens:

A tort is a species of wrong. A wrong is a breach of a duty owed to someone else. A breach of a duty owed to someone else is an infringement of a right they have against the tortfeasor... The law of torts is concerned with the secondary obligations generated by the infringement of primary rights. The infringement of rights, not the infliction of loss, is the gist of the law of torts.²⁶

²⁵ See Nicholas McBride, "Rights and the Basis of Tort Law" in *Rights and Private Law*, Eds., Nolan and Robertson, Oxford: Hart Publishing, (2012), pp. 331 – 367.

²⁶ Robert Stevens, *Torts and Rights*, Oxford: Oxford University Press, (2007), p. 2. See also Robert Stevens, "The Conflict of Rights", in *The Goals of Private Law*, Eds., Robertson and Tang, Oxford: Hart Publishing, (2009).

There are at least four types of rights employed by this analysis of private law, founded on Wesley Hohfeld's famous taxonomy.²⁷ The first is a claim right. A has a claim right that B *x* (or not-*x*), if and only if B owes A a duty to *x* (or not-*x*). Examples include the right not to be falsely imprisoned or the right to be provided with a competent professional service for which one has paid. A claim right possessed by A is directed towards some person or group of people, which means that the latter have a duty correlated to A's right. The breach of this duty is therefore not just a wrong *period*, but a wrong *to A*. The second class of rights are the privileges. A has a privilege to *x* if and only if A has no duty not to *x*. Examples of privileges include rambling on public rights of way or looking at someone on a bus. Possessing a privilege to *x* does not entail that others have a duty to allow one to *x*. A has no duty not to look at B on the bus, but equally B has no duty to allow A to look at her: she may move to where she cannot be seen if she wishes. The final groups are 'meta-rights'.²⁸ powers to alter the rights of oneself and others and immunities from the powers of others to alter one's rights.

Donal Nolan and Andrew Robertson identify four key features of rights-based analysis.²⁹ First, it is *structural* in the sense that it seeks to identify formal features of private law rather than its content. For instance, a rights theorist might emphasise the formal features that are common to rights in negligence, defamation and nuisance law, rather than describing the circumstances in which these rights are violated. Second, it is *monistic* in its central claim that private law can be best understood with reference to the concept of a right. This is usually thought to involve the rejection of competing models of explanation such as

²⁷ For the foundational analysis of legal rights, see Wesley Hohfeld, "Some Fundamental Legal Conceptions as Applied in Judicial Reasoning", *Yale Law Journal*, 23, (1913), 28 – 39, p. 32.

²⁸ See Francis Kamm, "Rights", in *The Oxford Handbook of Jurisprudence and Philosophy of Law*, Eds., Jules Coleman and Scott Shapiro, Oxford: Oxford University Press, (2002), p.479.

²⁹ See Donal Nolan and Andrew Robertson, "Rights and Private Law" in *Rights and Private Law*, Eds., Nolan and Robertson, Oxford: Hart Publishing, (2012), pp. 4 – 7.

policy-based reasoning.³⁰ Third, it is *formalist*, meaning that “the general, structural concepts of private law... determine the result (or the rule) to be applied in particular (types of) cases”.³¹ The formalism and monism of rights analysis are closely related. The formalist nature of private law means that decisions are predictable according to the correct application of principle, rather than particular policy determinations, which excludes the possibility of a deeply pluralistic understanding of private law. Finally, and most importantly for our purposes, rights-based theories are *interpretive*. They aim to provide a coherent and appealing explanation of private law as it stands, rather than a normative justification or critique.³²

In *Torts and Rights*, a definitive text in the recent surge of rights-based theorising, Stevens commits the vast majority of the book to demonstrating that rights theory explains the doctrines of private law better than the alternative ‘loss compensation’ model, which holds that the main objective of tort law is to specify when one party is liable to compensate for the loss of another. He argues that the loss compensation model cannot explain torts that are actionable *per se* (without proof of loss), such as libel and battery, and cannot account for remedies other than awards of compensation, such as injunctions and punitive damages. More positively, he contends that the rights model can provide a satisfactory explanation for torts that centrally involve compensation, such as negligence and slander, as well as the rules of limited recovery in pure economic loss and psychiatric harm.

³⁰ For mixed theories, this does not always involve a complete rejection of policy-based reasoning. Nicholas McBride argues that it would be irresponsible for the courts to disregard the public interest when deciding what rights should be recognised. See Nicholas McBride, “Rights and the Basis of Tort Law”.

³¹ Martin Stone, “Formalism”, in *The Oxford Handbook of Jurisprudence and Philosophy of Law*, Jules Coleman and Scott Shapiro Eds., Oxford: Oxford University Press, (2002), p. 189.

³² For example, see Stevens, *Torts and Rights*, pp. 4 – 5 and Allan Beever, *Rediscovering the Law of Negligence*, Oxford: Hart Publishing, (2007), p. 21.

Stevens recognises that the rights-based conception is essentially empty – it does not tell us what rights ought to be legally recognised and why. At the end of the book, he offers a rather half-hearted attempt to show why we have the rights conferred on us by private law.³³ He claims that we have moral rights, some of which ought to be given the force of law, and these “are capable of being deduced from the nature and experience of ourselves, and the world and society in which we live”.³⁴ The starting point for determining the content of these rights is the negative formulation of the Golden Rule, from which rights such as bodily integrity and free movement can be derived. This rule states that we should not do unto others what we would not want done to ourselves. This requires, Stevens thinks, that we “respect the choices and preferences of others”.³⁵

Stevens’ attempt to give this theory a solid normative foundation has already run into trouble. The Golden Rule, as he defines it, does not require us to respect the preferences of others, but to treat others as if they had our own preferences. Usually, this will prohibit serious offences such as murder, theft and so on, since almost all people prefer not to be subject to these acts; but not always. For example, one problem with Stevens’ view is that it seems to justify a right to suicide bombing. The suicide bomber causes death by explosion to others – a burden she also imposes on herself. In doing so, she demonstrably does not impose on others what she is not prepared to suffer herself. It is unlikely that Stevens favours suicide bombing at all, let alone a *legal right* to suicide bombing. This implies that the prospective suicide bomber could obtain a court injunction to ensure that the police do not interfere with her plans.

³³ He admits that a comprehensive account might call for a much longer treatment, but also insists that “there is no necessity for an extensive discussion of the works of Kant and Hegel in order to understand and justify the rights recognized by the common law”. See Stevens, *Torts and Rights*, p. 329.

³⁴ *Ibid.* p. 330.

³⁵ *Ibid.* p. 332.

Even when the Golden Rule has plausible implications about moral permissibility, it is unable to determine when impermissible acts should also be breaches of legal duties. Suppose that A prefers not to be negligently defamed. The Golden Rule implies that it is wrong for A to negligently defame others. But A might still prefer to live in a society where free speech is well protected and negligent defamation is not actionable in tort. She might accept, for example, that if she is negligently defamed by B, she ought not to have a cause of action against B, even though she is the victim of a wrong. The Golden Rule only holds that it is *morally* wrong for A to negligently defame others. It does not settle the matter about whether A should have a legal claim. But Stevens acknowledges that only some moral rights are legally protected and not others, so he owes us an explanation of what rights have the force of law and why.

Leaving this problem aside, The Golden Rule is also inconclusive in a different way. Suppose that A and B agree on matters of moral permissibility but disagree on matters of legal duty. A believes that actual malice should be required for a legal action, while B thinks negligence is sufficient. The preferences of A and B are inconsistent, and this poses a general problem. The Golden Rule does not tell us whether the law should follow the views of A or B, since both of them satisfy the criterion of offering others the same rights that they claim for themselves.³⁶ Needless to say, this is not a peripheral or eccentric counterexample, but goes to the heart of liability for reputational damage.

Stevens goes on to explain that although certain rights have good consequences they do not rely on these consequences for their justification. On the other hand, rights are not absolute and unqualified, and may be overridden when they come into conflict with other rights, or when greater values are at stake.³⁷ But we have now regressed to identifying the formal features of rights rather than providing a justification for them. These comments help

³⁶ For further critical discussion of the Golden Rule, see Derek Parfit, *On What Matters*, Vol. 1, Oxford: Oxford University Press, (2011), pp. 328 – 30.

³⁷ Stevens, *Torts and Rights*, pp. 335 – 7.

to clarify Stevens' view but they do not present any arguments in favour of it. There is little else in the chapter that offers a justification for the rights of private law, and ultimately he does not overcome the objection that rights-based theories are normatively shallow. In a work that engages in an intricate analysis of the content of private law, we are left with a vague gesture towards the Golden Rule as a reason for why any of it is justified.³⁸

C. Civil Recourse Theory

In a series of articles, John Goldberg and Benjamin Zipursky develop a new theory that holds that the purpose of tort law is to confer on victims of wrongs an avenue of civil recourse against their injurers.³⁹ Like Stevens, Goldberg and Zipursky seek to refocus tort scholarship on wrongs rather than the allocation of losses. They criticise a number of scholars, including Oliver Wendell Holmes, Patrick Atiyah, William Prosser and economic theorists such as Guido Calabresi, for conceiving of tort law as a method of addressing the

³⁸ Another account of when the law ought to recognise that A has a coercive right against B is given by Nicholas McBride. See Nicholas McBride, "Rights and the Basis of Tort Law", pp. 352 – 355. Roughly, his view is that A should be granted a coercive right that B do x if A has an interest in B doing x that is clearly sufficiently important to justify burdening B with the duty, and does not have side effects adverse enough to make it undesirable to grant A such a right. McBride's view is too underspecified to be satisfactory, but there is no space to defend that judgement here.

³⁹ See Benjamin Zipursky, "Rights, Wrongs, and Recourse in the Law of Torts." *Vanderbilt Law Review*, 51, (1998), 1–100; "Civil Recourse, Not Corrective Justice," *Georgetown Law Journal*, 91, (2003), 695–755; "A Theory of Punitive Damages," *Texas Law Review*, 84, (2005), 105–72; "Philosophy of Private Law" in *The Oxford Handbook of Jurisprudence and Philosophy of Law*, Eds., Jules Coleman and Scott Shapiro, Oxford: Oxford University Press, (2002), 623 – 655; and John Goldberg and Benjamin Zipursky, "Torts as Wrongs," *Texas Law Review*, 88, (2010), 917–86.

social problem of accident costs.⁴⁰ Against this conception, they reaffirm tort law as a coherent collection of wrongs, each involving “an interference with one of a set of individual interests that are significant enough aspects of a person’s wellbeing to warrant the imposition of a duty on others not to interfere with the interest in certain ways”.⁴¹ Like rights-based theories, civil recourse emphasises that torts are relational wrongs.⁴² Unlike crimes or regulatory infractions, which are wrong *tout court*, duties in tort are directives not to treat others in certain ways, violation of which constitutes a wrong to the beneficiary of the duty.

Goldberg and Zipursky argue that the loss allocation model cannot explain the central features of tort law. Like rights-based theorists, they question its ability to explain torts that are actionable *per se* or the diversity of available remedies. But vindicating a model of torts as a body of wrongs rather than a system of loss allocation is not their novel step – the unique move is the claim that tort law provides an avenue of civil recourse to the victims of private wrongs. A natural response to suffering a wrong is a unilateral act of aggression or reprisal against the wrongdoer. Since such reactions are prohibited in modern societies, the state is under an obligation to provide victims with some means of civil response against those who wronged them. It empowers claimants to co-opt the apparatus of the civil law to seek an appropriate remedy from the defendant. This is encapsulated in the right of action against the defendant afforded to the claimant. This right of action does not correlate with a duty on behalf of the defendant (that would be a primary right) but is instead a power to seek assistance from the court to enforce the claimant’s demands, thus altering the legal status of the wrongdoer.

Although Goldberg and Zipursky are hostile towards instrumentalism, their theory can in fact be cast in a conditional, unconditional or non-instrumentalist light. If the sole

⁴⁰ Goldberg and Zipursky, “Torts as Wrongs”, pp. 920 – 4.

⁴¹ *Ibid.* p. 937.

⁴² See Zipursky, “Rights, Wrongs, and Recourse in the Law of Torts”, pp. 59 – 63.

purpose of civil recourse is to satisfy a desire for revenge amongst the victims of wrongdoing that would otherwise spill out of control, and no constraints are placed on this goal, the theory is unconditionally instrumentalist. If its sole purpose is to fulfil a duty that the state incurs towards its citizens by claiming a monopoly on force, it is non-instrumentalist. If it is a combination of these purposes, whereby the attempts of the law to placate the desire for revenge amongst citizens are morally constrained, it is conditionally instrumentalist.

This ambiguity aside, civil recourse theory has a number of advantages. First, it is plausible that when a state prohibits aggressive responses to wrongs, it incurs an obligation to provide victims with a means of peaceful recourse. Since compensation schemes do not allow claimants to bring defendants to court, and criminal convictions are brought at the discretion of the state rather than the victim, a private right of action seems well suited to this role. Second, the concept of a right of action unifies different fields of private law, such as tort, property and contract, in which such a right is granted. It also distinguishes private from public law, as in the latter the state initiates an action and in the former it responds to an action brought by a claimant. Finally, it explains why some torts do not require proof of loss and some remedies are not compensatory. To take an example that has recently sparked debate, vindictory damages are easily explicable on civil recourse theory.⁴³

Goldberg and Zipursky present the theory as interpretive rather than normative. They leave it open whether such a system is justified. Defending civil recourse theory on moral grounds is a considerable feat. One would have to argue for a number of controversial propositions: that the criminal law is insufficient to satisfy a victim's desire for recourse against the wrongdoer; that victims' desire to harm defendants in response to their wrongs should be given institutional permission and support, even though many civil wrongs, such

⁴³ See *Ashley v Chief Constable of Sussex* [2008] UKHL 25 and *R. (on the application of Lumba) v Secretary of State for the Home Department* [2011] UKSC 12

as negligence, are relatively trivial; that the need for civil recourse outweighs considerations in favour of alternative systems, such as the wider availability of compensation under state-run schemes, and so on.

Moreover, civil recourse faces the same problem as rights-based and corrective justice theories: it does not specify in any detail when courts should impose remedial duties on defendants. What is required is either a defence of the grounds of liability that currently exist in private law, or an agenda for reform. Only when this is provided will we know whether the structural features of private law identified by Goldberg and Zipursky serve a defensible purpose.

D. Concluding Remarks

At the beginning of this chapter, I introduced two distinctions. The first was between interpretive and normative theory, and the second was the tripartite distinction between unconditionally instrumentalist, conditionally instrumentalist and non-instrumentalist theories. There is a clear division in the theoretical literature between unconditional instrumentalism and its detractors. The economic approach has dominated the instrumentalist camp. Many alternative views set themselves up in opposition to the economic theory and make considerable efforts to refute it.

Tort theory has also been dominated by an interpretive rather than normative approach. This approach has not been unsuccessful: many theories yield insights into the way tort law actually functions. Some of them make different but complementary points and can be combined with one another. John Gardner argues that, despite Goldberg and Zipursky's attempt to distance civil recourse from corrective justice theories, the two are

consistent.⁴⁴ Tort law does indeed give claimants a means of recourse, but one that is primarily corrective rather than punitive in character. Both of these approaches are also consistent with rights-based theories. Weinrib explicitly uses the language of rights as well as corrective justice and it is no coincidence that Stevens' arguments against the loss compensation model parallel those of Goldberg and Zipursky. Tracing the differences and similarities between these theories is not my main purpose, however, so I will say no more about it here.

As I have emphasised throughout, some of these views do not purport to be normative at all and some only partially normative. Of those that do not entirely eschew the justificatory enterprise, there is a degree of slippage between explanatory and normative modes of theorising. Epstein and Stevens' work is affected by the circularity that results from this slippage: they rely on the fact that their theories are consistent with case law in order to give them normative support, the purpose of which should be to show that they are justified independently of their fit with case law.

Perhaps the most important omission is the general failure to address the grounds of liability in tort. Most of the theories have little to say about what set of primary rights and duties ought to be recognised by the law.⁴⁵ Posner, Epstein, Weinrib and Stevens all suggest principles from which these might be derived (wealth maximisation, libertarian rights, Kantian abstract agency and the Golden Rule respectively), but none of these are developed

⁴⁴ See John Gardner "Torts and other Wrongs", *Florida State University Law Review*, (2012), (forthcoming).

⁴⁵ This omission is most problematic from a normative perspective, but it can also be an objection for explanatory theories. For example, Gregory Keating argues that corrective justice fails to recognise that tort is a law of wrongs, not just a law of redress for wrongs. The remedial obligations that arise from the breach of primary duties are therefore an important but subordinate aspect of tort law. See Gregory Keating, "The Priority of Respect over Repair", *Legal Theory*, 18, (2012), 293 – 337.

or defended in any detail. In Posner, Weinrib and Stevens' work they are subordinated to the main aim of explaining tort law, and Epstein fails to deliver a complete account of the libertarian rights on which his view depends.

Other scholars have noted the need for a normative approach to tort law. Nicholas McBride compares the fixation of private law theorists on interpretive questions with the fruits of the Lotus plant in Homer's *The Odyssey*. According to the story, on their homecoming to Ithaca after the Trojan War, Odysseus sends some of his men to explore an island containing Lotus fruit. Anyone who eats the fruit loses interest in doing anything else, and the more men Odysseus sends, the more fail to return. McBride's suggestion is this: like the Lotus fruit, interpretive puzzles beguile theorists, distracting them from the equally important normative question, 'what basic rights should tort law give us against other people?'⁴⁶

3. Normative Tort Theory and Tort Law

Some theories that we have discussed, such as civil recourse and right theories, are applicable to private law generally – duties in private law include those under contract, unjust enrichment and equity as well as tort. Other theorists, such as Jules Coleman (whose work we will explore in the next chapter), develop a specific theory of tort as a component of a broader understanding of private law. It is important for interpretive theories to be clear about the content of the law and the definitions of legal categories, as this is the primary data to be processed and explained. This is less pertinent for a normative theory, which is not beholden to these technical legal facts. Nevertheless, it will be useful to describe in greater detail the relationship between normative tort theory and the law.

⁴⁶ Nicholas McBride, "Thinking about Tort Law – Where Do We Go from Here?", (unpublished ms), pp. 9 – 10.

There is a prior question here. Before we can describe this relationship, we need a definition of tort law. This is a notoriously difficult task and has been pursued as a scholarly end in itself. Again, unpacking this question reveals a cluster of others. How is tort law different from other areas of private law, such as contract? How does private law differ from public law? On the latter question, private and civil wrongs are often taken to be synonymous. Of course, this depends on how we define them. One definition is that a private or civil wrong is a breach of a legal duty imposed for the benefit of a private party, rather than the public at large. If A breaches such a duty owed to B, she has not just committed a wrong, but a wrong in relation to B. Another definition, which we have also noted in relation to civil recourse theory, is that it is a wrong capable of affording the victim, or the victim's estate, a civil right of action against the wrongdoer.⁴⁷

As Goldberg and Zipursky argue, the conferral of a civil right of action helps to distinguish private law from criminal law and most of public law. Categorising the subclasses of private law has proven a little more difficult, however. Some traditional views on the division between them have fallen out of favour. For example, Percy Winfield held that duties in tort are primarily fixed by law, whereas duties in contract are fixed by the parties themselves.⁴⁸ This view can be criticised from two directions. It falls short as a

⁴⁷ Both definitions are mentioned by Peter Birks. See Birks, "The Concept of a Civil Wrong" in *Philosophical Foundations of Tort Law*, David Owen Ed., Oxford: Oxford University Press, (1995), pp. 31 – 51, at pp. 33 and 40. Rights-based theories generally emphasise the relational duty feature of private law, whilst civil recourse theory emphasises that private wrongs generate civil actions. McBride and Bagshaw note that the two definitions come apart. If A negligently injures B in the course of a robbery, A has breached a relational duty to B but any civil action will be cancelled by the illegality defence. See Nicholas McBride and Roderick Bagshaw, *Tort Law*, 4th Ed., Pearson, (2012), pp. 9 – 10.

⁴⁸ Percy Winfield, *The Province of the Law of Tort*, Cambridge: Cambridge University Press, (1931), p. 32.

definition of contract because it ties liability too closely to the making of promises. It cannot allow contractual liability to follow the conferral of a benefit by the claimant or reasonable reliance on the defendant's behaviour. It is also inadequate as a description of tort because a number of obligations in tort law are, in a sense, fixed by the parties themselves. This is true of liability for careless misstatements that defendants undertake voluntarily.⁴⁹

Some have sought to salvage a version on Winfield's thesis. McBride and Bagshaw argue that tort law specifies those basic rights we enjoy free of charge and without having to make special arrangements. They underplay the significance of 'assumption of responsibility' as a basis for tort liability, since rights arising in this way are not paid for – their creation is just specific to certain circumstances.⁵⁰ Contract law gives us a facility to alter our tort law rights, either by gaining more rights or reducing the rights we have against others, all of which must be 'paid for' in some sense. The resort to contract as a way of securing extra rights is partially undermined by the expansion of liability for assumptions of responsibility, but McBride and Bagshaw contend that, in general, the distinction is coherent and decisive.

Similar disagreements abound about the relationship between torts and equity, property, unjust enrichment, crime and public law. Tort law, at least as it is perceived and practiced, is a heterogeneous creature that fosters enough examples and counterexamples to fuel these debates. Consider another dispute about the relationship between tort and public law. Public nuisance is sometimes considered part of tort law, but is not civilly actionable without proof of special damage. This seems to undermine the requirement that a civil wrong confers a private right of action. But some have simply argued that public nuisance is

⁴⁹ For a general discussion, see Simon Deakin, Angus Johnston, and Basil Markesinis, *Markesinis and Deakin's Tort Law*, 6th Ed., Oxford: Clarendon Press, (2008), pp. 17 – 18.

⁵⁰ McBride and Bagshaw, *Tort Law*, pp.17 – 18.

not a tort at all, and has unfortunately come to be considered so by a series of accidents.⁵¹ Equally, products liability, vicarious liability and the rule in *Rylands v Fletcher* are believed to be constituents of tort – they are certainly taught to students in courses entitled ‘Tort Law’ – but they do not involve liability for *wrongs*. Again, some have argued that these sources of liability are not properly part of tort law.⁵² Of course, we might incorporate them into tort law by conceptualising them as forms of faultless wrongdoing, but strict liability undeniably puts pressure on the view of tort as a law of wrongs.

These debates are important for a number of reasons. As John Gardner argues, the legal designation of wrongs has enough practical implications to make it prudent to coherently separate torts from other civil wrongs. Each body of law also tends to develop its own way of justifying its existence, and collapsing the boundaries is likely to create an incomprehensible mess rather than a unified whole.⁵³ But at the normative level, these debates are not very significant. The finer points of legal classification do not have much bearing on what kind of claims we ought to recognise, although it might affect how we should categorise them, and perhaps give us clues about how best to justify them.

Although I will continue to use the label ‘normative tort theory’, I recognise that, depending on one’s favoured definition, some of the issues I discuss are outside the ambit of tort law. If tort law is a law of wrongs, then – at least on one interpretation of wrongdoing – my discussion of liability for harm caused by permissible action falls outside its scope. Perhaps it is better to begin by specifying the ethical questions that seem most relevant and important, before asking whether they come under the rubric of tort law. A core issue I address is the moral basis of liability to compensate in cases meeting the following

⁵¹ See John Spencer, “Public Nuisance – A Critical Examination”, *Cambridge Law Journal*, 48, (1989), 55 – 84.

⁵² See McBride and Bagshaw, *Tort Law*, pp. 13 – 16.

⁵³ John Gardner, “Torts and Other Wrongs”, *Florida State University Law Review*, (2012), (forthcoming).

conditions: (1) A causes harm to B, (2) the harm is unintentional, and (3) the *ex poste* losses cannot be spread amongst the population, but must be borne by A, B or both of them. We can further divide these cases into two types, those where A's actions are wrongful (negligence) and those where they are not (strict liability), and I discuss both of them. If strict liability is properly part of tort law, then both are paradigmatic tort cases. But if it is not, this part of the discussion is not strictly a question for tort theory.

There is another important difference worth noting: in real cases the losses can usually be shifted to others through some loss spreading mechanism, such as insurance or loss compensation schemes. Indeed, some of the theoretical literature argues that tort should be partially or completely replaced by state-run insurance schemes.⁵⁴ These alternatives to tort law must be evaluated at some point down the line, but it is best to establish the correct principles in simpler cases first. Only then will we know what to compare with loss spreading considerations.

There are many aspects of tort liability that I will not discuss. I will omit liability to compensate for intentional torts, such as battery, as well as defamation, nuisance and many others. I will also generally restrict the discussion to compensation. Torts that are actionable *per se* and alternative remedies such as vindicatory or gain-based damages raise important questions, but my primary interest is in harm and its correction. That said, my discussion will also develop an analysis of corrective duties independently of compensatory damages. This is because I think that the most philosophically perspicacious account of corrective duties does not map accurately onto the legal categories as they have emerged. I distinguish, for example, between negating and counterbalancing duties, and note that both can be fulfilled by a payment of compensation. This is one example of the general methodological

⁵⁴ There are also those who argue that tort law should be abolished, leaving the private insurance market to deal with the cost of accidents. See Peter Cane, *Atiyah's Accidents, Compensation and the Law*, 7th ed., Cambridge: Cambridge University Press, (2006).

approach taken by this thesis, which maintains that it is best to explore the most fundamental theoretical questions independently of legal doctrines, categories and principles.

Methodology and Outline

This thesis will be largely concerned with the definition and significance of corrective duties. Very broadly speaking, this might place it within the corrective justice camp. I am a little reluctant to apply this label, as although there are similarities between the present project and corrective justice approaches, there are also important differences that must be acknowledged. In order to illuminate the issues, I will first discuss some corrective justice theorists and then lay out the structure of the thesis in greater detail. Corrective justice is an ancient idea, dating back to Aristotle's *Ethics*, but the modern theory was developed by Jules Coleman and Ernest Weinrib, with whom we will begin. I will then consider John Gardner's contribution, which is the most convincing account of corrective justice and its role in tort law. A key difference between corrective justice and the present work is that I am only interested in normative issues. This means that a very different methodology is apt for the project and I will defend this methodology in section 3.

1. Corrective Justice

What is the theory of corrective justice? This is not a straightforward question for three reasons. First, the term 'corrective justice' is rather general and invites misinterpretation. If it is defined as offering an appropriate remedy in response to a wrong, corrective justice is purely formal and has no implications about the content of private law rights. Indeed, this is a common criticism levelled against it.¹ Admittedly, some theorists attach a substantive doctrine of private law to these formal features, but the term 'corrective justice' is unable to distinguish between different views. Second, it is not clear whether corrective justice should

¹ See Peter Cane, "Corrective Justice and Correlativity in Private Law", *Oxford Journal of Legal Studies*, 16(3), (1996), 471 – 488.

be restricted to wrongdoing. People are harmed by many other events besides wrongs, and if these were corrected by a state-run compensation scheme, it is hard to see why such a scheme would not be implementing corrective justice. As we will see, most theorists circumscribe the concept by limiting it to relational wrongs and duties. Finally, there are differences between the two best known proponents of the theory, Jules Coleman and Ernest Weinrib, which make it difficult to elucidate a single, coherent view. But despite these differences, some core similarities can be revealed. These common threads distinguish Coleman and Weinrib from those we considered in the previous chapter.

A. Jules Coleman

Coleman's thinking on corrective justice has changed in two central ways. In his earlier writing he claimed that corrective justice provides a complete account of tort law.² In later work he adopted a mixed theory, repudiating the idea that the practice can be understood as a unified whole.³ Contract law, Coleman later argued, can be explained on economic grounds. It facilitates mutually beneficial transactions by reducing uncertainty about potential compliance with agreements.⁴ Some aspects of tort law – what he calls the 'penumbra' – are also explained by economic efficiency. A court may impose liability on a manufacturer whose product carries a sub-optimal warning, even if that warning did not cause the claimant's injury. This cannot be an instance of corrective justice, since a causal link is necessary before such a duty is owed – the ground of liability is instead the encouragement of more efficient warnings.⁵ Nevertheless, the core of tort law implements

² See Jules Coleman, *The Practice of Principle*, Oxford: Oxford University Press, (2001), p. 34.

³ See Jules Coleman, *Risks and Wrongs*, Oxford: Oxford University Press, (2003).

⁴ *Ibid.* part 2.

⁵ *Ibid.* p. 387.

corrective justice. Intentional and unintentional torts fall into the core whilst no-fault schemes, enterprise liability doctrines, workers' compensation and most of product liability fall on the periphery.

The second change is in the conception of corrective justice itself. In earlier writing Coleman defended the annulment thesis.⁶ According to this thesis, the point of corrective justice is to eliminate, rectify or annul wrongful losses. In *Risks and Wrongs*, Coleman rejects this view because it supplies only the grounds of recovery rather than the mode of rectification. When A wrongs B, it may be true that B has suffered a wrongful loss that should be compensated, but the annulment thesis does not explain why A should compensate B. B's wrongful loss could just as easily be compensated by a third party or a state-run scheme, and so the annulment view does not capture the structure of a private action, where the injurer incurs a special duty to pay compensation once a successful claim is brought by the victim.

To fix this defect, Coleman looks to the relational account of corrective justice. The relational view establishes "a scheme of rights and responsibilities between individuals... unlike the annulment thesis, it creates specific, agent-relative reasons for acting."⁷ But although Coleman accepts that the best conception of corrective justice generates agent-relative duties of repair, he does not wholly accept the relational view. The relational view holds that wrongs ought to be corrected by imposing a duty on the wrongdoer, but it does not accept that the loss gives content to the duty. Its emphasis is on the wrong one does rather than the magnitude of loss resulting from the wrong. It should follow from this, Coleman says, that two parties who commit the same wrong but cause different degrees of harm should incur the same duties. After all, if their wrongs are identical and corrective

⁶ See Jules Coleman, "Tort Law and the Demands of Corrective Justice", *Indiana Law Journal*, (1992), 349 – 379.

⁷ Jules Coleman, *Risks and Wrongs*, p. 315.

justice compels them to repair the wrong, rather than the consequence of their behaviour, then the content of their duties should also be identical.

Having dismissed this result as inconsistent with legal practice, he is left with a mixed conception of corrective justice that combines aspects of the annulment and relational views, involving a duty to repair wrongs and wrongful losses. In his words: “The duty to repair the wrong flows from the relational view; the importance of wrongful losses to the demands of corrective justice is the remnant of the annulment view.”⁸ Coleman makes a final refinement, however. He removes the implication of the relational view that wrongdoers must correct their wrongs as opposed to wrongful losses. This is for two reasons. First, other principles, such as retributive justice, already provide for the correction of wrongs. Second, some wrongs in tort law are not defeasible by excuse and therefore may not reflect a moral fault in the agent. These points suggest that we should abandon the view that corrective justice seeks to repair wrongs as such. Rather, it seeks to repair wrongful losses. But this is not a reversion to the annulment thesis, since the agent-relative duty of repair is retained from the relational view: “corrective justice imposes on wrongdoers the duty to repair wrongful losses their conduct occasions, losses for which they are responsible.”⁹ The insight taken from the relational view is that duties under corrective justice are agent-relative. The insight of the annulment view is that the content of the duty is determined by the magnitude of the claimant’s loss.

B. Ernest Weinrib

Weinrib defends a version of the relational account of corrective justice, which he embeds in a more general formalist theory of private law. According to Weinrib, the normative source

⁸ *Ibid.* p. 324.

⁹ *Ibid.* p. 325.

of the rights and duties recognised by private law is the Kantian concept of abstract agency, which he claims is “the normative grounding for private law’s governance of particular harms”.¹⁰ Kant’s ‘principle of right’, from which more specific provisions are derived, permits the free will to realise itself consistently with a similar right for others. In practice, this involves a right to one’s bodily integrity, which houses the will, and some external objects such as personal property. Although the basic rights and duties of private law arise from the structure and significance of free will, welfare (conceived in terms of wellbeing) plays a secondary role, since these rights “crystallize certain welfare advantages and protect them from wrongful interference”.¹¹ But rights rather than welfare have primacy; private law rectifies setbacks to welfare only insofar as this is required to correct a wrongful interference with a right.

To explain this corrective function, Weinrib invokes the Aristotelian conception of corrective justice that involves the rectification of wrongful gains and losses. Although the most common remedy available to claimants is an award of damages, the gains and losses corrected by a system of private law are essentially normative rather than factual. So what is a normative loss? Weinrib explains that corrective justice embodies norms that set the terms of fair interaction between people. Action that conforms to a relevant norm maintains the normative equality between the parties, but action which violates it necessarily produces a gain for the injurer and a loss to the victim.¹² A normative loss, it turns out, consists in having one’s rights violated, and a normative gain consists in violating the rights of another.

Although Weinrib, unlike Coleman, seeks to derive the content of private law rights (in very broad outline) from an abstract notion of agency, the core of Weinrib’s theory is not dissimilar to Coleman’s. Perhaps the most significant difference is Coleman’s refinement of

¹⁰ Ernest Weinrib, *The Idea of Private Law*, Cambridge, MASS: Harvard University Press, (1995), p.129.

¹¹ *Ibid.* p. 131.

¹² *Ibid.* pp. 115 – 117.

the relational view. Weinrib is comfortable with the assumption that correcting a wrong sometimes involves compensating for the losses stemming from that wrong. But Coleman insists that this is a misunderstanding of what it means to correct a wrong. He refers to cases where wrongdoing benefits rather than harms the victim, such as a taxi driver whose recklessness causes a person to miss her flight, which subsequently crashes. In these cases, the duty to correct the wrong cannot be conflated with the duty to undo the consequences of a wrong. This would yield the absurd result that the taxi driver has a duty to kill her passenger. Coleman thus recommends that we retain the distinction between the wrong and the wrongful loss, with corrective duties directed towards the latter.

This disagreement aside, Coleman and Weinrib's views share four central features. First, the duties under corrective justice are agent-relative. When A commits a wrong against B, A incurs a duty to perform some corrective action with respect to B. A has special reasons to fulfil this action that do not apply to others because only A has committed the wrong. Second, this duty is relational. It is owed to B and imposed on A for B's benefit, rather than the benefit of all. On a strong interpretation of this feature, A's duty cannot even be owed to a general pool of victims of which B is a member, but only to B herself. Third, responsibility is a condition of incurring a corrective duty. Coleman uses the term 'responsibility' rather than 'wrongdoing' since he believes that one can be liable to compensate for harm one has caused even if one acted justifiably. But both thinkers reject the idea that third parties or the state can perform corrective justice by compensating for losses – only the person who is responsible for the harm bears the duty (although Coleman believes that the operation of corrective justice is conditional on the absence of other institutions that might serve similar functions).¹³ Finally, in standard cases, the content of the duty is determined by the degree of harm done. The defendant must, so far as possible, pay a sum of compensation commensurate with the magnitude of harm she inflicts.

¹³ Coleman, *Risks and Wrongs*, p. 402.

C. The Defence of Corrective Justice

What arguments do these theorists offer in support of corrective justice? Coleman and Weinrib both defend their views, at least in part, by criticising those that they seek to replace. Coleman attacks aspects of the economic theory¹⁴ and in other work offers a brief critical discussion of George Fletcher's nonreciprocity view.¹⁵ Coleman explicitly adopts a mixed methodology. His primary purpose is to explain tort law, but in doing so he reveals the moral importance of corrective justice. I will say no more about this here since the appropriate methodology for normative tort theory is an important question and will be discussed in section 3.

Weinrib is less clear about how he intends to defend his theory. He is very critical of legal instrumentalism, the view that tort law is comprehended and justified according to the goals it seeks to pursue. Although he does not make the distinction, it is likely that he would reject it in both its conditional and unconditional forms. His anti-instrumentalism is so thorough that he even denies that private law serves to implement the *moral* norm of corrective justice. According to Weinrib, there would be no moral norm of corrective justice were it not already instantiated in a juridical system of duties.¹⁶ Interestingly, this move seems to rob him of the most promising way of defending his theory against its competitors: to argue that it is more effective than others in realising the moral norm of corrective justice and that this norm is independently justified. Instead, Weinrib only offers cryptic comments about the justification of his formalist theory.

There appear to be two ways in which he believes his theory derives its normative force. The first is its internal coherence and the second is its relationship to the Kantian

¹⁴ See Jules Coleman, *Markets, Morals and the Law*, Oxford: Oxford University Press, (1998), part 2.

¹⁵ Coleman, *Risks and Wrongs*, pp. 252 – 269.

¹⁶ Weinrib, *The Idea of Private Law*, p. 110.

notion of abstract agency. Weinrib claims that the private law relationship is “a mode of moral association that attaches decisive importance to the justification of the norms that constitute it. Coherence is important for private law relationships because it is indispensable to justification.”¹⁷ Weinrib proceeds to explain the function of coherence:

The point of adducing a justification is to allow that authority to justify whatever falls within its scope. Thus if a justification is to function as a justification, it must be permitted, as it were, to expand into the space that it naturally fills. Consequently, a justification sets its own limit. For an extrinsic factor to cut the justification short is normatively arbitrary.

These comments are rather mysterious. Weinrib follows up with an example designed to clarify the matter:

Consider... the effect of the causation requirement on the operation of the insurance rationale. Once we accept the plausibility of allowing injured plaintiffs to tap into the defendant’s liability insurance, the defendant, if insured, should be held liable regardless of his or her role in causing the injury... From the standpoint of the insurance rationale, the defendant’s causal role is an arbitrary factor, no more relevant a reason for limiting loss-spreading than is the color of the defendant’s hair.¹⁸

Judging by these passages, it seems that Weinrib simply intends to say the following: a justification for some practice, X, must justify all components of that practice otherwise it is not justifying X but some other possible practice.

¹⁷ Ibid.

¹⁸ Ibid.

This is conceptually true, but it provides no normative support for Weinrib's theory. If our purpose is to justify all and only those principles and practices that constitute private law, then Weinrib is right that we will fail to do this if the justification we offer only supports some components and not others. For example, if our overriding argument is that all victims of accidents ought to be compensated, we might end up justifying a system that replaces tort law, such as a no-fault compensation scheme. I will later argue that our aim should not be to justify private law as it currently exists anyway – this would assume a specific set of answers to the questions that a moral inquiry must ask. But even if one's purpose was to justify private law, all Weinrib is offering here is a methodological constraint.

Perhaps coherence should be understood as implicitly entailing certain forms of justification.¹⁹ There are two versions of this interpretation. On the restrictive version, the justification must be of the right kind. A loss compensation scheme might be internally consistent, but would not be coherent in the relevant way as it does not embody the right kind of justification. The right kind of justification is one that centres on the correlativity between the claimant and defendant, understood in the Weinribian sense. On the less restrictive interpretation, coherence involves a unity of justification or an absence of pluralism. Private law becomes incoherent, then, when it mixes correlativity with another form of justification such as loss spreading. What makes it incoherent, on this view, is not the introduction of the wrong kind of justification, but the combination of two mutually exclusive forms of justification.

The restrictive interpretation is a tendentious understanding of coherence, but without it it is hard to see how Weinrib can reject alternatives such as loss compensation models, as these may be coherent in the relevant sense, even if they embody a different form of justification. Leaving this aside, however, there is still an element of mystery. Not only do

¹⁹ This interpretation seems to be favoured by some commentators. See Sandy Steel, "Private Law and Justice", *The Oxford Journal of Legal Studies*, (forthcoming).

we have no explanation of why correlativity provides the best kind of coherence, so far we know little about the normative attraction of correlativity itself.

Weinrib's reliance on Kantian abstract agency may help us here. There is intuitive appeal in a system of law that allows each person's free will to realise itself in the manner of that individual's choosing, compatible with a similar freedom for others. The importance of correlativity derives from the relationship between the defendant and claimant that is created when a rights violation occurs. Weinrib dedicates a chapter to explicating 'Kantian right', but unfortunately his discussion is rather abstract. The only substantive conclusion he draws about the content of private law rights is that they should include bodily integrity and some external property. Apart from this, the account is largely empty, leaving most of the important questions unanswered. In later work, Weinrib defends some other conclusions. He states that Kantian agency leads us to accept no liability for omissions, restricted negligence liability for pure economic loss, extensive property rights and a system of redistributive taxation that provides for only the most basic means of sustenance for the poor.²⁰ But the derivation of these more specific conclusions from the concept of agency requires a more detailed argument than Weinrib supplies.

There is another reason why we should doubt Weinrib's ability to show us that a particular set of rights can be derived from the formal features of free will. The plausibility of any proposed bundle of rights depends on its implications across a range of cases, but these implications cannot be assessed by Weinrib's methodology. To do this, we need to test fundamental principles against hypothetical (or real) cases. Adopting this method also ensures that our principles are not overly vague; that is, they yield a sufficient number of implications in concrete cases to be informative and action-guiding. There is a suspicion that Weinrib's appeal to abstract agency cannot yield the requisite degree of specificity. For example, why does Kantian agency allow for restricted liability for pure economic loss rather than rejecting it altogether? The answer seems to be simply that this is the current

²⁰ Ernest Weinrib, *Corrective Justice*, Oxford: Oxford University Press, (2012), Ch. 8.

position in tort law. But there was a time, before *Hedley Byrne v Heller* and its progeny, when such liability was non-existent. Is this previous state of affairs somehow in deep conflict with Kantian agency? How exactly do the fundamental constituents of Kantian moral philosophy support the decision in *Hedley Byrne*? Until Weinrib offers a convincing answer, we have reason to worry that his starting point is too vague.

D. John Gardner

Coleman and Weinrib have developed the most detailed theories of corrective justice, but some important work has recently been done by John Gardner, which in certain respects improves on Coleman and Weinrib, so it is worth briefly noting these advances to set the stage for the arguments about corrective duties advanced in the remainder of the thesis.

Gardner recognises that Coleman and Weinrib are ambiguous about whether they offer a normative, interpretive or mixed defence of corrective justice (he also thinks that there is no bright line between justifications and explanations since to justify is also to offer a certain kind of explanation).²¹ The lack of attention paid by Coleman and Weinrib to the *moral* norm of corrective justice leads to a significant point. Weinrib thinks that the moral norm is fully constituted by the legal one, and Coleman holds more modestly that the moral norm is partially constituted by the legal one. For Weinrib the moral norm simply does not exist beyond the legal one, and for Coleman the legal content of corrective justice sharpens the more indeterminate moral norm. Gardner responds that even if either of these constitutive relationships is exemplified by tort law, the law must still be evaluated as an instrument. Perhaps not an instrument for wealth maximisation as the economic theorists would maintain, but at least an instrument for the implementation of the moral norm. After

²¹ See John Gardner, "What is Tort Law For? Part 1. The Place of Corrective Justice", *Law and Philosophy*, 30, (2011), 1 – 50, p.3.

all, if enforcing the legal norm is *self-defeating*, the law has failed. But it could only be viewed as a failure if we accept that it fails instrumentally to satisfy the underlying moral norm. As a parallel, consider counter-productive traffic regulations that make the roads more dangerous. One who believes in a moral norm of safe driving can point out that the regulations are self-defeating because they make the roads more dangerous. One who believes that the moral norm of safe driving is constituted by the legal norm cannot make sense of the idea that the regulations are self-defeating, as there is no independent standard of success by which the regulations can be measured.²²

Gardner's insight inspires a broader theoretical point. If we accept Weinrib's strong constitutive claim, it is hard to see how any justification can be given for corrective justice at all. Weinrib insists that duties of corrective justice exist *given* a system of private law, but then what reason do we have to create or maintain this system? For all Weinrib argues to the contrary, we might be better off without it. If we abolished private law tomorrow, there would no longer be any duties of corrective justice left unfulfilled, and we would save a lot of time and money too. In fact, there are far more unfulfilled duties under an active system of private law, since many rights violations either do not make it to court or are unfairly denied. Given this, would we be better off abolishing private law to reduce the amount of unfulfilled corrective duties? The obvious reply to this suggestion is that there are *moral* corrective duties, and far more of these would be left unfulfilled if we abolished private law.

Gardner offers two other insights worth stating. The first is that the best account of corrective justice is pluralistic in the sense that it operates to force wrongdoers to take corrective measures *and* to deter wrongdoing. Although he does not use the terminology, this probably means that it is a conditionally instrumentalist theory, which permits deterrence as a legitimate aim on the condition that only those who already owe corrective duties can be used for this end. Anti-instrumentalists such as Weinrib reject the deterrent

²² *Ibid.* p. 19.

role of tort law, but this role is not inconsistent with corrective justice. In fact, as Gardner observes, it is performed simply by *doing* corrective justice.

Secondly, Gardner offers the most perspicacious explanation of why corrective duties should follow the breach of primary duties. He explains this in roughly the following way. We have a number of reasons that go in to establishing a primary duty. When this duty is breached, many of the reasons that support the original duty are still operative, and these reasons might support a new duty. When we attend to the kind of reasons that survive the breach of a primary duty, we find that they add up to create a corrective duty. In chapter 4, when we analyse corrective actions in greater detail, we will make use of Gardner's 'continuity thesis'. In chapter 5, when we investigate the moral basis of corrective duties, we will give content to the continuity thesis by specifying exactly what kind of reasons go in determining the core corrective duties.

2. Development of the Thesis

In this section I will give a more detailed summary of the thesis. A foundational question for any theory in the corrective justice tradition is this: in respect of *what* are corrective duties owed? For an interpretive theorist, the answer is relatively straightforward: corrective duties are owed in respect of those forms of damage or loss that are legally recognised. But what is it to which *moral* corrective duties respond? Although I do not assume a single answer to this question (indeed, I argue to the contrary), it is reasonable to think that these duties mainly correct *harm*. In chapter 3 I therefore develop a theory of harm that helps us understand the core corrective duties. Such a theory enables us to characterise harm, thus understanding its moral importance, and to identify and measure harm.

I canvass a number of different views. I consider whether harm should be defined as setback to preferences, interests or wellbeing, and conclude that the wellbeing view is most defensible. The preference and interest-based views are too broad and too narrow: too broad

in that they falsely pick out some non-harmful events as harmful and too narrow as they fail to identify some harmful events. The flaws of these views reveal the strength of the wellbeing view.

In order to properly identify harmful events and come to conclusions about the correct measure of harm, I address the problem of pre-emption. I argue that pre-emption cases are best dealt with by a *complex* version of counterfactualism. The *simple* counterfactual view is familiar to lawyers. It holds that an event harms a person if she is worse off than she would now be had the event not occurred. The simple view does not yield plausible results in pre-emption cases where, had an event, E, not made V worse off, another event, E2, would have made her as badly off or worse off. The simple view counterintuitively implies that E does not harm V. The best way to deal with these cases is to allow our theory to pick out different counterfactuals when making judgements about the same case, or so I argue.

With this theory in hand, I proceed to the second fundamental question that a theory of corrective justice should answer: what actions count as corrective? In particular, I draw a distinction between negating and counterbalancing. For clarity, I restrict my discussion of these actions to harm. I define negating as rendering a victim's future wellbeing *identical* to what it would otherwise be and counterbalancing as rendering a victim's future wellbeing *equal* to what it would otherwise be. I show how the failure to recognise this distinction causes confusion in discussions of corrective duties and obscures important normative questions. I then argue that there is a morally significant difference between the two forms of corrective action that has implications for moral permissibility. I explain this difference in terms of the distinctive wrongness of interference. Interfering with the distribution and content of a person's wellbeing is distinctively wrong, even if that person's wellbeing is not reduced overall.

In chapter 5 I address a philosophical problem of actionability. Actionability concerns the question of whether the consequences of wrongdoing, or non-wrongful conduct that occasions liability, are compensable or subject to other corrective actions. Even if it is

established that A wrongs B, it is a further question whether the consequences of the wrong are actionable. Harm that is too remotely connected with the wrong is not actionable, and neither is pure economic loss in most circumstances. Legal limitations on actionable loss may be justified by practical constraints. Trivial harms are not actionable because they are too small to justify the litigation costs involved in ascertaining and quantifying them. But at the moral level there are also important questions about actionability. I will focus on one in particular. I am interested in how we should respond to a victim's false beliefs about her own wellbeing. The two most important types of cases are the following: (1) those in which a victim is benefitted but she mistakenly believes she is harmed (or not benefitted) and (2) those in which a victim is harmed but she mistakenly believes she is benefitted (or not harmed). I focus on this issue partly due to constraints of space, but also because it has not been discussed in the literature, despite its similarity to debates in liberal political philosophy.

I explore three views. *The Objective View* holds that compensation should always be determined according to the correct conception of harm and the victim's false beliefs disregarded. *The Respect View* holds that compensation should always be determined according to the victim's sincerely held beliefs about wellbeing and the objective conception disregarded. *The Mixed View* holds that the victim's beliefs should be followed in (1)-type cases and the correction conception of wellbeing should be followed in (2)-type cases, so compensation should be offered in both cases. *The Mixed View* also claims that where a victim suffers actual harm, compensation should be quantified according to the correct account of wellbeing, but when a victim does not suffer harm or is benefitted, any claim she possesses, grounded in respect for her autonomy, is limited to a reasonable threshold. I then defend *The Mixed View* against its competitors.

In chapter 6 I investigate the moral basis of corrective duties in cases that meet the following conditions: (1) A causes harm to B, (2) the harm is unintentional, and (3) the *ex poste* losses cannot be spread amongst the population, but must be borne by A, B or both of them. I consider George Fletcher's theory of reciprocity and Gregory Keating's revisions,

but ultimately reject them. In doing so, I defend a set of alternative principles that yield more plausible results in core cases. These principles are grounded in the avoidability of harm, the benefits generated by the harm-causing activity, and considerations of distributive fairness. To borrow Gardner's framework, these three principles give us the main reasons that add up to make a corrective duty following harmful conduct. Finally, in chapter 7, I explore the relationship between corrective and distributive justice. I undermine the arguments of some theorists that corrective justice is independent of distributive justice. I unpack the ways in which these forms of justice might be independent and reject these claims. I also show that distributive concerns can affect the moral basis of liability, depending on whether loss spreading is possible. This argument extends and compliments the earlier claim that considerations of distributive fairness partly determine when corrective duties are incurred. This chapter also presents the opportunity to ask whether corrective duties, as a distinct deontological category, have any fundamental significance. I argue that they do not (a little incongruently given the language of corrective duties adopted throughout the thesis) and that it is more logical and enlightening to focus on the broader distinction between agent-relative and agent-neutral duties rather than corrective and non-corrective duties.

3. Normative Methodology

This thesis adopts a purely normative methodology. Although I will occasionally refer to case law, the success of the claims I defend does not depend on their interpretive capacity, but rather on their ethical plausibility. Part of what makes a claim ethically plausible is its match with *intuitions* about *cases*.²³ An intuition is a considered, non-inferential judgement.

²³ Two refinements must be made to this quick summary. First, most theorists are not only interested in cases, but work back and forth between cases and principles to achieve what John Rawls called reflective equilibrium. Second, an intuition about a case is not a decisive argument.

It is considered in the sense that it is arrived at after a reflective process and need not be an unstable gut-reaction. It is also non-inferential as it is a direct response to the case, rather than a conclusion reached by inferring the implications of a principle to which one has a prior commitment. Since intuitions are non-inferential, they are not merely a product of our pre-existing theoretical commitments²⁴ and give us a way of testing principles. A *case* is a situation in which an ethical judgement is required. These judgments might be about whether some act is permissible, whether an agent is blameworthy, who is liable to bear the costs of certain actions, and so on. Cases can be real or hypothetical, although as I will argue later hypothetical cases give us the best analytic tool with which to reach sophisticated normative conclusions.

There is another methodology, outlined most explicitly by Coleman, which synthesises explanatory and justificatory frameworks. It might be argued that this methodology suffices for reaching normative conclusions, and it has the added benefit of revealing to us the moral principles inherent in legal practice. Before we can legitimately employ a purely normative methodology, therefore, Coleman's view must be addressed.

Coleman eschews presenting his defence of corrective justice as a purely normative endeavour. In *The Practice of Principle*, he states:

Intuitions usually have to be supported by consistence with other intuitions or an elucidation of the moral idea behind the intuition.

²⁴ It may be that our favoured principles shape our intuitions. If so, there is something mischievous in the claim that intuitions are non-inferential. They may be non-inferential, but they are still determined by our principles because they are conditioned by them. It is unlikely, however, that intuitions are always or fully determined in this way. If this were true, we would not be able to show that a principle to which we are committed has a counterintuitive implication, but this is a common enough phenomenon in philosophical argument.

In arguing that tort law is best explained by corrective justice I do not mean to be defending tort law thereby. As it happens, the conception of corrective justice embodied in tort law expresses important moral values, and in what follows I will say something about these values and the ways in which corrective justice expresses them. Still, even if tort law is best explained by corrective justice *and* corrective justice is an important and independent moral ideal, it does not yet follow that tort law represents a justified—let alone a morally required—institution.²⁵

Coleman conceives of his methodology as an exercise in ‘middle level’ rather than ‘top down’ theorising. Top down theory proceeds from a set of political or moral principles, thought to exert moral force, to a model of justified institutional structures against which actual institutions can be measured.²⁶ Middle level theory, on the other hand, begins with an analysis of the principles that are evident in the law as it is currently practiced. It explains significant elements of practice as the embodiment of principle, and only then asks how attractive these principles are. The justificatory project is not prior to the explanatory one, but an outgrowth of it.

Middle level theory is one aspect of Coleman’s broader commitment to philosophical pragmatism. Coleman summarises his pragmatic method in terms of several basic characteristics, which include the following views: that the content of concepts is to be explicated in terms of their inferential role in the practices in which they figure; that sometimes a philosophical explanation of a practice takes the form of showing how certain principles are embodied in it; that the way in which a concept figures in one practice

²⁵ Jules Coleman, *The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory*, Oxford: Oxford University Press, (2001), p. 5.

²⁶ *Ibid.*

influences its proper application in all others, and so practices are to be viewed holistically; and a commitment to the in-principle revisability of all beliefs, categories of thought etc.²⁷

Coleman makes a few comments in elaboration of these claims. For example, the ‘inferential role semantics thesis’ is central to his interpretation of law. This thesis holds that the content of a concept can sometimes be analysed in terms of the practical inferences that can be drawn from its use. If Smith utters the phrase, ‘I promise to meet you for lunch today’, I can infer that Smith has established a duty to meet me for lunch. The concept of a promise is, in part, revealed by the practical inferences I draw when someone makes me a promise. Coleman’s claim about the interdependence of meaning can also be illustrated with reference to the institution of promising. Promises feature in a variety of different practices – personal, political and legal. Coleman’s suggestion is that the meaning of promising in one practice influences its meaning in others.

I will not offer an appraisal of Coleman’s broader pragmatist legal theory. He is right to emphasise that the law is a human-driven, practical enterprise, but demonstrating the truth or falsity of his core pragmatic claims would require more attention than I am able to give it here. Instead, I will focus on the integration of explanatory and justificatory modes in his methodology.

Coleman is not dismissive of top down (or purely normative) legal theory, but he does reject it as a candidate for his own work. Regrettably, he does not defend his preferred method against it, and we are left with only one throwaway comment as to why he rejects it. He states that the practical standards that issue from top down accounts “are generally at a level of abstraction that hovers above, and is consistent with, a broad range of different and non-compossible legal institutions.”²⁸

We might level a similar charge against Coleman. The principle of corrective justice states that “individuals who are responsible for the wrongful losses of others have a duty to

²⁷ *Ibid.* p. 6.

²⁸ *Ibid.* p. 5.

repair the losses”.²⁹ This principle is formal; it does not offer substantive accounts of ‘responsibility’ ‘wrongful loss’ or ‘reparation’. As such, it is consistent with a broad range of views about the doctrinal content of tort law. It does not tell us what interests should be protected and to what extent; how balances should be struck between competing values, such as liberty and security; whether strict liability is justified and why; what losses are actionable, and so on.³⁰ It is strange that the inability of a theory to yield concrete prescriptions about legal rights and duties should worry Coleman given that his own method is not well suited to do so either.

This is not a decisive objection, however. It is not a problem that his methodology does not derive detailed practical conclusions if it is not designed to seek them. The more important objection is that his doubts about normative theory are misplaced. For one thing, the best methodology is not solely top down, in Coleman’s terminology. It does not begin with indubitable moral principles and then deduce their consequences for the design of legal institutions. Instead, it is closer to John Rawls’ method of reflective equilibrium, in which the theorist works back and forth between basic moral principles and their implications in particular cases. Neither level of the theory is immune to revision. If a principle that seems plausible in the abstract produces a deeply counterintuitive implication in a concrete case, the principle may be revised. Equally, some of our intuitions about cases may need to be rejected if they conflict with principles that give intuitive results in other cases.³¹

Another motivation for the objection might be that the analysis of hypothetical cases cannot be easily transposed into practice. There are many practical concerns that would prevent the principles that best explain these intuitions being directly enshrined in the law. This is clearly true and no theorist need deny it. But it would be an exaggeration to conclude

²⁹ Ibid. p. 15.

³⁰ Coleman offers a slightly more substantive account in *Risks and Wrongs*.

³¹ For a critique of reflexive equilibrium, see Francis Kamm, *Morality, Mortality, Volume 2, Rights, Duties and Status*, Oxford: Oxford University Press, (1996), Introduction.

that normative theory is incapable of generating implications for real cases. First of all, this has not proven true in other contexts. To draw another parallel with the philosophy of criminal law, competing principles of criminalisation frequently indicate how the criminal law should be reformed, even though they operate at an abstract theoretical level.³²

The mere existence of a practical impediment rarely shows that we should abandon a principle. Perhaps the clearest cases in which it does so are those where legal enforcement would be self-defeating. For example, even if there is a moral duty of easy rescue, recognising a legal duty may encourage people to avoid placing themselves in situations where they have to perform rescues. This might reduce the overall amount of rescues, thus defeating the purpose of the law. On other occasions, however, the situation is more complex. The legal enforcement of moral principles is not self-defeating, but there are other practical issues. Perhaps the law is too blunt an instrument to fully realise the relevant principle; perhaps enforcement of the law will be difficult or expensive, and so on. In these circumstances, we must weigh the benefit of legally recognising the principle against the relevant costs. In order to do this, we need to know what the best principles are, independently of practical constraints. Only then can we decide if the costs involved in their legal enforcement outweigh the value of enforcing them. Consequently, normative analysis seems a reasonable place to begin our enquiries.

Perhaps Coleman fails to appreciate the potential for normative theory to draw concrete conclusions about tort law because he overlooks the role of hypothetical case analysis. Studying hypothetical scenarios in which competing principles yield conflicting results allows us to test principles in a very fine-grained range of cases. A rigid focus on

³² See, for example, Joel Feinberg, *The Moral Limits of the Criminal Law, Vol. 1, Harm to Others*, Oxford: Oxford University Press (1987); *The Moral Limits of the Criminal Law, Vol. 2, Offence to Others*, Oxford: Oxford University Press, (1988); Arthur Ripstein, "Beyond the Harm Principle", *Philosophy and Public Affairs*, 34, (2006), 215 – 45.

current law, by contrast, may blind us to important cases, testing the limits of our principles, which have not arisen in the development of the common law. Contrary to Coleman's suspicions, case analysis seems well equipped to give specific guidance for the obvious reason that it is based upon intuitive judgements about specific cases. Notwithstanding the caveat that judgements in hypothetical cases will not always be transferrable (or fully transferrable) to legal practice, this method is a direct way of making principled arguments about how we should construct legal doctrines.

Moreover, there may be more than one principle that plausibly explains a doctrine in tort law, and if so we have reason to adopt that which best explains intuitions in hypothetical cases. If we can, we should seek to find principles that yield plausible results in the widest possible range of cases. It might be objected that the selection of co-extensive principles has no practical relevance. If two different principles give the same answers in all cases that actually arise, there is no reason to adjudicate between them. But usually we cannot know this in advance. It is difficult to rule out the possibility that future cases might reveal a gap in our analysis, and we are better off adopting a policy by which we thoroughly explore whatever cases test our principles, including the unlikely and the bizarre. We can also defend a stronger claim. Suppose we could identify the precise set of cases, past and future, that would engage two principles, and discovered that the principles always give identical results. We would still have reason to determine which should be preferred by testing them against hypothetical cases, assuming an answer was available. By making this enquiry, we wouldn't obtain further guidance about how we ought to act, but we might learn more about our moral landscape. We might find out that it is more pluralistic or monistic than we thought; that our moral ideas have certain relations we had not predicted, and so on. Although the first task of normative theory is to tell us what we ought to do, non-practical conclusions such as these are not to be dismissed as vacuous or irrelevant.

Finally, this methodology also gives us an analytic tool that we can use to separate and examine distinct and sometimes conflicting considerations that are otherwise liable to be conflated. Consider two factors in a standard case of unintentional harm: A voluntarily

chooses to engage in an activity that could foreseeably harm B and A makes this decision for his own benefit. These are distinct factors and have been conflated by some theorists, as I will explain in chapter 6. Using case analysis we can tease them apart and test them against each other.

There is more to be said in defence of the case-based method, and more criticisms of it that must be refuted. For example, it is sometimes thought that hypothetical cases are too schematic to tell us anything important about our moral duties. The study of these cases through comparison and variation assumes the existence of principles that apply to multiple cases that share some relevant feature, even if they differ in other respects. Moral particularists doubt that such principles exist: the determination of our moral duties is too sensitive to the contextual richness of each individual case for such analysis to be useful.³³ Others note that our moral intuitions about hypothetical cases can be affected by irrelevant factors, such as the way the questions are framed or our hidden biases, and conclude on this basis that the method is deeply flawed.³⁴ Yet others think that judgements about bizarre or fantastical cases cannot tell us anything important about morality in the real world.³⁵ I do not think that any of these criticisms undermine the case-based method, although some give

³³ See, for example, Jonathan Dancy, *Ethics without Principles*, Oxford: Clarendon Press, (2004).

³⁴ For an account of the research in experimental philosophy that could be used to support this claim, see Daniel Kahneman, "The Cognitive Psychology of Consequences and Moral Intuition", *The Tanner Lecture in Human Values*, University of Michigan, Ann Arbor, (November 1994). For criticism, see Francis Kamm, *Intricate Ethics: Rights, Responsibilities, and Permissible Harms*, Oxford: Oxford University Press, (2007).

³⁵ See Allen Wood, "Humanity as an End in Itself", commentary in Derek Parfit, *On What Matters*, Vol. 2, Oxford: Oxford University Press, (2011), pp. 66 – 82.

useful insights about how best to apply it. Unfortunately there is no space here to give them the rejection they deserve, so I will say no more about them.³⁶

Instead, I will offer a brief critique of Coleman's alternative methodology. Fletcher seems also to adopt this method by analysing landmark cases and doctrines, seeking to uncover their unifying rationale. Fletcher clearly believes that the paradigm of reciprocity not only explains the development of tort law, but also represents an attractive principled basis for the law. Given the work of these scholars, is a purely normative methodology really necessary? Can middle theory, as well as showing us the principles immanent in tort practices, provide a justification for them?

As Coleman puts it, "justificatory questions are not prior to the explanatory ones, but can grow out of the explanatory project as it reveals the abstract principles in greater specificity and concreteness".³⁷ One problem with this is that if these principles are to be justified thoroughly, they must be defended against other candidates. But then would it not be simpler to ask which of the principles that purport to be the moral basis of tort liability is preferable? If the 'outgrowth' that Coleman describes is to be treated as a fully grown rather than a stunted limb of the theory, then the normative questions are surely separable from the explanatory ones.

Uncovering principles that are immanent in tort practice is not even guaranteed to reveal the best possible justification of that practice. As I mentioned previously, more than one principle may explain elements of the current law. This may be true of principles that are otherwise quite different. Epstein's theory of causation and fault-based theories will

³⁶ It is also unnecessary to repeat the work of others who have made important arguments in defence of the method. See Victor Tadros, "Responses", *Law and Philosophy* 32 (2-3), (2013), 241-325, pp. 245 – 7 and *The Ends of Harm: The Moral Foundations of Criminal Law*, Oxford: Oxford University Press, (2011), pp. 6 – 8; Jeff McMahan, "Moral Intuition" in *The Blackwell Guide to Ethical Theory*, Ed., Hugh LaFollette, Blackwell Publishing, (2000).

³⁷ Coleman, *The Practice Principle*, p. 6.

converge on cases where both fault and causation are present. A large amount of tort cases fall within this overlap. But without assessing cases where A causes harm without fault and those where A is at fault but causes no harm, we will not know which principle provides the most robust support for the core cases. This is a striking example, but I suspect that other principles whose distinction is more nuanced will also overlap in this way.

The relationship between explanation and justification is also unclear on Coleman's account. Explanations and justifications are not simply correct or incorrect, compelling or indefensible – they possess these properties to varying degrees. Suppose we must choose between two principles, A and B. A is better at explaining the current law than B, but seems less justifiable, while B is more justifiable but a poorer explanation. Which is to be preferred? The obvious answer is that A is preferable as an explanation and B is preferable as a justification. Together, these conclusions might then give us reason to reform the law. If this case is a potential quandary for the theorist (and I am inclined to believe that it is quite common) then explanations and justifications ought to be conducted separately.

So our tentative conclusion is that a mixed methodology cannot fulfil the central *desiderata* of normative theory. It cannot tell us how best to justify our current practices, since the principles that actually underlie them may not be the best available. In any case, as I mentioned in the discussion of Weinrib, our primary question should not be 'what is the best justification for tort law?' This will preclude the perfectly real possibility that the best principles justify some other set of social and political institutions that might take the place of tort law. Instead, we should return to first principles. What are the *best* answers to the cluster of ethical questions for which tort law provides *one* set of answers? In formulating our methodology in this way, we take a great step away from the tort scholars we have discussed, by removing our focus from tort law as it is presently practiced and defined. Discussion of the common law itself is the old territory of tort theory. By entering a new territory – the realm of the fundamental ethical questions that the common law seeks to answer – we will find a way of re-charting the old.

Harm

Harm is a fundamental concept in many normative claims: its significance is certainly broader than tort law. For example, in the political context, it is sometimes argued that the only justification for the use of state coercion is the prevention of harm to others, or that it is impermissible to forcibly prevent someone harming themselves. The application of the concept is broader still, as harm is relevant to many ethical judgements about permissibility. But it is also a logical place to begin our study of tort theory, since it gives shape to our most important corrective duties: duties to compensate. Compensatory duties do not *only* arise in response to harm,¹ but it is fair to assume that they typically do. Moreover, the measure of compensation owed in these cases is generally determined by the magnitude of harm inflicted. In order to identify and quantify compensation, therefore, we must develop an adequate theory of harm.

What kind of concept is harm? Does it contain normative as well as non-normative elements? Normative judgements are closely tied to facts about harm, but ultimately our theory should make reference only to non-normative facts. In other words, it should be capable of being administered by someone with no knowledge of what acts are permissible or impermissible. This is because normative judgements often turn on facts about harm. If facts about harm depended on normative judgements, this would introduce circularity into our thinking. Whether A harms B should not depend on whether A acts permissibly, because whether A acts permissibly often depends on whether she harms B.

At this point, we might raise a sceptical worry. Is the term 'harm' really substantive? Does it refer to a distinct property, or is any disagreement about its meaning merely terminological? One way of thinking about harm is as a shorthand term for a set of

¹ I will not have space to deal with this thoroughly, but I will later suggest some circumstances in which compensatory duties are not only owed in respect of harm. See chapter 5.

permissibility-affecting facts. ‘Harm’ is a term of art that is stipulated to cover those non-normative facts that go into determining whether some act is permissible. But this approach faces a problem. Either harm covers all such non-normative facts or a subset of them. The former conclusion makes an unwarranted assumption. There may be other general concepts that are importantly distinct from harm that unify subsets of permissibility-affecting facts (such as non-harmful interference). On the other hand, if we leave this possibility open by adopting the latter view (taking harm to cover a subset of permissibility-affecting facts) we implicitly appeal to some criteria that distinguish harm from other such facts. It seems plausible that harm is one among a number of concepts that refer to a related set of permissibility-affecting facts, so it is worth investigating our intuitions about putative cases of harm to see if they can be systematised into a coherent theory.

Moreover, we could run a similar sceptical argument about causation. We could raise doubts about whether the phrase ‘A causes B’ refers to any non-stipulative relation. There are many well-entrenched disagreements in the philosophy of causation and it is not certain whether our intuitions will ever support a coherent, non-stipulative theory. Despite this parity, it is difficult to abandon the idea that causation is something real. At this stage, at least, it is reasonable to persist with the search. Even if the attempt to develop a non-stipulative theory of harm fails, the many cases considered along the way may reveal new insights and problems that assist our understanding of the subject matter around which discussions of ‘harm’ centre.

There are two questions that a theory of harm ought to answer. First, what is the currency of harm? That is, if harm is setback to X, what is X? Harm may consist in a setback to the victim’s powers, resources, will, wellbeing, some other unspecified currency, or even a plurality of currencies. Settling the currency issue will *characterise* harm. It will tell us what kind of good or goods setback to which constitutes harm. It will also help to *identify* harms, since it will exclude setback to those goods which do not form part of the relevant currency. Second, what is a *setback* to X? For example, if wellbeing is the correct currency, a victim’s wellbeing could be setback by an event, E, in a number of different ways. E could

setback V's wellbeing if she is worse off than she would now be had E not occurred (the simple counterfactual view); or worse off now than she was prior to E (the simple temporal view); or just badly off, as a result of E, according to some non-comparative measure (the non-comparative view). A theory must adjudicate between these, and perhaps other, possibilities. Settling the setback issue will further enable us to identify harm, particularly in cases where the harm suffered has multiple causes or is overdetermined. It will also allow us to calculate the magnitude of harm and therefore judge the measure of compensation to which a victim is entitled.

The two key questions are separable. It is possible to ask what constitutes a setback to the currency, leaving 'currency' as a placeholder for whatever view on the former issue turns out to be correct. They are equally important, however, and I will address them in turn. I will begin by arguing that the currency of harm is wellbeing. This view will be given a two-part defence. First I will offer a potted critique of two of its competitors, which specify preferences and interests respectively as the appropriate currency. These views are too broad in the sense that they label some non-harmful events as harmful and too narrow as they fail to identify some harmful events (similar arguments can be made about the way they characterise benefits).² I will then give a positive argument in favour of wellbeing as the currency of harm.

In the second part of the chapter, I address the setback issue. I do not discuss all views in the comparativist family: I omit temporalism, which has been criticised elsewhere.³ I argue instead that we ought to accept a complex counterfactual view. The simple view holds that we should always compare the victim's actual state with the state she would now

² Victor Tadros criticises Seana Shiffrin and Robert Nozick's views on similar grounds. See *What Might Have Been*, (forthcoming).

³ For an argument that the only plausible version of temporalism relies on counterfactual comparisons, and that insofar as temporalism is distinct from counterfactualism it is false, see Victor Tadros, *What Might Have Been*.

be in had the event not occurred. The complex view is distinct from this in that there is no single counterfactual state to compare with the victim's actual state in all cases. There are insurmountable problems with the simple version of counterfactualism, but the solution is not to switch to one of the competing views, as there are insurmountable problems with these too. Instead, the solution is to buy plausibility at the cost of simplicity and develop a more sophisticated version of the counterfactual view.⁴

1. The Currency of Harm

In *Anarchy, State and Utopia*, Robert Nozick offers a theory of compensation based on the hypothetical preferences of the victim. He argues that “something fully compensates a person for a loss if and only if it makes him no worse off than he otherwise would have been; it compensates person X for person Y's action A if X is no worse receiving it, Y having done A, than X would have been without receiving it if Y had not done A. (In the terminology of economists, something compensates X for Y's act if receiving it leaves X on at least as high an indifference curve as he would have been on, without it, had Y not so acted.)”⁵ Although Nozick's is an analysis of compensation, we could also hold a preference-based view of harm, according to which a person is harmed if and only if her preferences are frustrated (and benefitted if and only if her preferences are satisfied).

A similar view is developed by Seana Shiffrin, who argues that “typically, harm involves the imposition of a state or condition that directly or indirectly obstructs, prevents,

⁴ Tadros' central thesis that we ought to accept some version of the complex counterfactual view is also broadly similar to that of the present chapter. In order to defend this thesis fully, however, the challenge posed to counterfactualism by Matthew Hanser's event-based account must also be addressed, and competing versions of the complex view refuted.

⁵ Robert Nozick, *Anarchy, State, and Utopia*, Wiley-Blackwell: New Ed., (2001), p. 57.

frustrates, or undoes an agent's cognizant interaction with her circumstances and her efforts to fashion a life within them that is distinctively and authentically hers—as more than merely that which must be watched, marked, endured or undergone.”⁶ This passage admits of various interpretations. One plausible reading is that harm is defined as setback to an agent's attempt to realise wilful objectives. By wilful objectives, I mean those objectives that relevantly emanate from an agent's will, and are not the product of cravings or manipulation. In this respect, Shiffrin's view is an improvement on the preference account, which does not (in its current form) place any conditions on the preferences that determine harms and benefits. Rather than exploring the differences between these two views further, I will outline some general problems faced by both of them.

One difficulty is that they imply that certain beings – those that can experience physical pain but do not possess wills or the capacity to form preferences – cannot be harmed. Non-human animals, the young and some mentally impaired adults do not possess wills, but can experience physical pain and can therefore be harmed. Of course, this depends on exactly how we understand the will or preference-formation. If we understand these things in a behaviouristic sense, so that something wills a state of affairs if, under the right conditions, it would attempt to bring it about, then many non-human animals possess wills. Nevertheless, inactive beings that are receptive to pain, such as comatose humans, do not possess wills or preference-forming capacities. I take it as self-evident that the experience of physical pain is sometimes harmful. I refrain from a universal claim here as sensations of pain can also be part of a complex pleasure. But any theory of harm is inadequate if it implies that these types of beings are *incapable* of being harmed.

The harmfulness of physical pain poses other problems. Even in those cases involving pain that *are* harmful on these views (when pain conflicts with an agent's will or preferences) they tend to mischaracterise the harm involved. If harm just is a setback to

⁶ Seana Shiffrin, “Wrongful Life, Procreative Responsibility, and the Significance of Harm”, *Legal Theory*, 5, (1999), 117 – 148, pp. 123 – 4.

wilful objectives, then we are unable to explain why intense pain is typically contrary to an agent's will by reference to its harmfulness. Similarly, if harm is setback to preference-satisfaction, then we cannot explain why people usually prefer not to be in pain. Again, this seems to get things backwards. Physical pain is contrary to an agent's will because it is intrinsically harmful; it is not harmful because it is contrary to the will. Both views lack a capacity that a theory of harm should plausibly possess: that of explaining why agents often prefer to avoid harmful states. If this is correct, then the harmfulness of experiencing pain must be independent of its effect on a person's will or preference structure. Hence these views not only fail to correctly identify cases of harm, but mischaracterise harm generally.

Besides mischaracterising harm, both theories also have implausible implications in many cases. Consider the following:

Kidney Donor: Pete decides to donate one of his kidneys to Joe, a homeless person who lives in his neighbourhood. Although Joe faces suffering and death, he has a strong concern for Pete's wellbeing and prefers Pete to keep both his kidneys. But Pete insists and undergoes a painful operation to donate his kidney, after which he makes a number of lifestyle changes to adjust to life with one kidney.

In *Kidney Donor*, the preference view implies that Pete is benefitted and Joe is harmed. This gets things backwards again: the intuitive verdict is that Pete is harmed by giving away his kidney and Joe is benefitted. These are all-things-considered judgements, of course – Pete enjoys some benefit by doing a good deed and Joe incurs some harm by seeing Pete suffer, but these are outweighed by the respective harm and benefit of losing and gaining a kidney. Pete and Joe's preferences are also relevantly connected to their will, so Shiffrin's view fares no better in *Kidney Donor*.

Kidney Donor raises a general problem, which is that harm is not always contrary to an agent's will or preferences. When inflicting pain, suffering or deprivation on someone does not conflict with their will or preferences, it is still sometimes appropriate to

characterise it as harm, but this is precluded by both views. This is most clearly seen in cases of self-sacrifice, where one willingly suffers harm for the sake of another goal. For example, I may throw myself in the path of a bullet that would otherwise hit my daughter, or even a bullet that would otherwise hit my beloved Mercedes. I may experience a sense of satisfaction that I have protected something I love, or furthered a valuable goal, but I have suffered harm nevertheless. We can sharpen the example by excluding such satisfaction if we suppose that, in order to realise a posthumous goal, I willingly suffer a painful death, the manner of which renders me unable to enjoy my personal sacrifice while I am alive. In such a case I unmistakably, and willingly, suffer harm. On the will and preference-based views all of this is incomprehensible – even an agonising death cannot constitute harm if it is willingly endured. This is clearly wrong: self-sacrifice is an intelligible phenomenon which often entails harm, so we should reject any view which cannot make sense of it.

In light of these criticisms, Shiffrin might reply that if Pete decides to suffer a loss of one kind for the sake of a benefit of another kind, he does not will the loss for its own sake. The loss merely furthers a more important aim, so there is nothing implausible about the claim that Pete's will is not frustrated all things considered. This reply does not show that the will theory has plausible implications in *Kidney Donor* after all – it bites the bullet. It is true that saving Joe's life means more to Pete than avoiding the harm of losing a kidney, but it does not follow that he is better off overall by donating. It is more natural to say that he harms himself for the sake of a goal he values more than avoiding harm. These views still imply that self-sacrifice, insofar as it involves harm to oneself, is conceptually impossible.

Let us now briefly consider the view, proposed by Joel Feinberg, that harm consists in setback to interests. He states that “there can be wrongs that are not harms on balance, but there are few wrongs that are not to some extent harms.”⁷ This is because most wrongs

⁷ Joel Feinberg, *Harm to Others: The Moral Limits of the Criminal Law*, New York: Oxford University Press, (1984), p. 35.

interfere with some interest. The invasion of a welfare interest is the most serious but not the only kind of harm a person can sustain: interference with non-welfare interests can also be harmful. By appealing to interests, Feinberg holds that a person can be harmed by an event even if it does not impact on their experiences. Conversely, not all negative experiences or sensations are harmful. Experiences can “distress, offend or irritate us, without harming any of our interests... These passing unpleasantnesses are neither in nor against one’s interests. For that reason, they are not to be classified as harms.”⁸ Feinberg draws a generic contrast between harmful conditions and unpleasant physical and mental states, which he subcategorises as hurts, offenses and ‘other’. His use of the term ‘unpleasant’ here is rather misleading since, on his view, physical ‘hurts’ can be very severe without constituting harm as long as they do not impede an interest.

Feinberg takes pains to draw a distinction that will strike many as implausible. When a hurt is sufficient to impede a person’s interests, as when pain distracts someone from their business, “it is the capacity of the condition to distract, rather than its inherent character as a hurt, that renders it a harm”.⁹ This means that the same set of unwanted physical sensations can harm A and not B if it interferes with a goal that is adopted by A but not B. There is a tension here, which Feinberg acknowledges, between the extent to which people have interests independently of their choices and the extent to which their interests are determined by their choices.

This tension can be highlighted by returning to cases of self-sacrifice. If we accept that genuine self-sacrifice is possible, does it follow that suffering harm can be in a person’s interests? Feinberg says that the “satisfaction of an immediate want is in one’s interest... only when it is a means to the promotion of more ulterior ends in which one has an interest”.¹⁰ More precisely, the question is this: can a choice to suffer harm ever represent the

⁸ *Ibid.* pp. 45 – 6.

⁹ *Ibid.* pp. 47 – 8.

¹⁰ *Ibid.* p. 56.

promotion of an end in which one has an interest? If a person's interests are shaped by their desires and choices, it seems that it can, given the determination some individuals show when self-sacrificing. Interests, moreover, are more detachable from the experiences of an agent than wellbeing, lending further support to the idea that it can be in an agent's interest to undergo a painful experience.

However, if harm just is setback to interests, then the concept of harm is in tension with the notion of furthering an interest through suffering. Feinberg could give a similar response to the one we imagined for Shiffrin, that self-sacrifice involves frustrating one interest for the sake of a different interest. We might say that in this case a non-welfare interest conflicts with a welfare interest and the non-welfare interest takes precedence. The agent therefore does not suffer harm because she acts according to her overall interests. But the fault in this reply is the same as before. Although it is reasonable to perceive a conflict of interests here, it is counterintuitive to claim that the agent suffers no overall harm, especially in cases of self-sacrifice involving death. Whatever goal is furthered, it is unlikely that this is more in the agent's interest than the continuation of her life. Even if the interest for which she sacrifices herself is sufficiently important, and the decision is rational and admirable, it seems perverse to think that a painful death that deprives the agent of a full and healthy life is not harmful.

In criticising these views, I have mainly focussed on the fact that they are too narrow: they do not encompass some harms such as those involved in self-sacrifice. It can also be argued that they are too broad.¹¹ The interest view is too broad because people have an interest in others not disrespecting or attempting to harm them, but these things are not intrinsically harmful. Disrespect may cause harm if it leads to insecurity or depression, and attempts can obviously harm people if they are successful, but the interest theory implausibly implies that these things are harmful in themselves. Turning to the preference view, consider the following:

¹¹ See Victor Tadros, *What Might Have Been*, p. 5.

Narcissus: Narcissus injures himself and becomes paraplegic. He takes comfort from the fact that his face is untouched. He is so vain that he prefers life without mobility to any permanent scarring on his face. After witnessing his injury, Zeus takes pity on him and restores the use of his limbs with a lightning bolt that leaves a faint scar on his face.

The preference view implies that Zeus harms Narcissus, but this seems implausible. Narcissus is better off with a faint scar on his face than he is in a state of permanent paralysis; it is just that his vanity prevents him from realising this. Perverse preferences pose problems because they can cause people to be mistaken about their own wellbeing. The preference view cannot capture the general truth that a person can prefer what is bad for them.

If these views are defective, what view should we accept? It is difficult to give a complete defence of the positive claim that harm is setback to wellbeing, but it is possible to raise a presumption in its favour. To do this, a few problems with the wellbeing view must be noted and dispelled. The first concern is that wellbeing is a deeply contested concept. How can we reasonably adopt this view without a comprehensive account of wellbeing? There are two related worries here. One is that the definition of harm will be vague without a full elaboration of 'wellbeing'. Given that separate theories will have different implications as to precisely what events setback an agent's wellbeing, there will be a group of events that are neither harmful nor non-harmful on the view defended here. The second worry is that quite apart from its vagueness, the definition will not be justified until one conception of wellbeing is defended against others.

These problems can be overcome. With respect to the first worry, it must be remembered that there is a subgroup of events that setback an agent's wellbeing on *all* plausible accounts. This is probably true of extreme physical pain. If we restrict the discussion, as I have done, to cases involving such events, the criticisms will be acceptable

to those who accept conflicting conceptions of wellbeing. Even in non-standard cases, the vagueness problem can be overcome by either adjudicating between accounts of wellbeing on an *ad hoc* basis, or simply hedging bets by noting that a particular conclusion depends on a prior question that remains unanswered. This latter possibility is something of which we ought to be aware, but it hardly renders a theory of harm impossible. The second worry sets the standards of justification too high. The initial task is to defend the definition of harm as setback to wellbeing against other possible currencies. Assuming wellbeing can be defended on independent grounds, the fact that a fully justified account of wellbeing has not been adduced will not impact upon this conclusion. And as I will now argue, there are positive grounds to prefer wellbeing to other currencies.

There are two general arguments for this conclusion. The first is that the inability of competing views to identify and characterise harm seems to stem from their failure to acknowledge the role of wellbeing in harming. Physical pain, in most cases, is a straightforward example of setback to a basic element of wellbeing: the absence of negative physical sensations. Now return to cases of self-sacrifice, in which the preference and wellbeing analyses conflict. The preference analysis implies that the agent is not harmed, despite the diminishment in wellbeing, because her preference is satisfied. The wellbeing analysis implies that the agent *is* harmed, despite the satisfaction of a preference, because of the diminishment in wellbeing. The intuitive conclusion is that in at least some cases of self-sacrifice, the agent is harmed, so the wellbeing analysis yields the right result. This gives us reason to suspect (and a similar argument can be run for the interest account) that the preference analysis fails to generate the right implication because it does not recognise the role of wellbeing in harming.

The second general argument is that issues relating to harm and wellbeing appear to be very similar, even when these terms have not been fully unpacked. It is no coincidence that the various theories of wellbeing and harm proposed by philosophers have similar form and content. Accounts of wellbeing are commonly divided into three types: experiential

theories, desire theories and ‘objective list’ theories,¹² and philosophers have had equivalent discussions about harm. The specific disagreements that philosophers have about wellbeing have natural correlates in debates about harm. So as the following parallels indicate, the contested nature of wellbeing counts in favour of this view.

In his discussion of death, Thomas Nagel considers whether harm can take a non-experiential form.¹³ After all, how can anything be bad for a person without affecting their conscious experiences? Experiential theories of wellbeing draw on the same thought in claiming that an agent’s wellbeing is dependent solely on the quality of her experiences. This division is evident in the debate about whether a person can be harmed by a posthumous event. There is disagreement, for example, about whether failing to implement a person’s will or destroying the projects they achieved while they were alive harms them.

There are also affinities between desire theories of wellbeing and the preference-based view of harm. In this context, desires and preferences are synonymous, so if one adopted both of these views, any setback to a preference would be both a reduction in wellbeing and a harm. I noted that Seana Shiffrin defines harm as (approximately) setback to an agent’s attempts to realise wilful objectives. Such a setback seems to involve a failure to satisfy a desire or preference. Suppose A breaks her leg and is no longer able to realise a variety of goals. On Shiffrin’s view, A is harmed because she has suffered an obstruction to her “cognizant interaction with her circumstances”. This involves the failure to satisfy her desire to fulfil these goals and therefore constitutes a setback to wellbeing on desire-based theories.

Finally, there are many possible theories of harm, corresponding to ‘objective list’ theories of wellbeing, which share the implication that harm involves setback to a good irrespective of the agent’s desire for that good. It might be argued that the rejection of

¹² See Derek Parfit, *Reasons and Persons*, Oxford: Clarendon Press, (1984), pp. 493 – 502.

¹³ See “Death” in Thomas Nagel, *Mortal Questions*, Cambridge: Cambridge University Press, (1979).

meaningful human relationships is harmful, even if it satisfies an agent's hermitic preferences. This is the idea we appealed to in *Narcissus*. If one believes that the loss of an objective good is bad for a person's wellbeing, even if she prefers the loss, it is natural to think that the loss is also harmful. These considerations are not decisive, but they raise a presumption in favour of wellbeing as the correct currency of harm, and this presumption is defensible independently of a comprehensive account of wellbeing.

2. Defining Setbacks

A. Criticisms of the Counterfactual View

I have argued that harm is setback to wellbeing, but there are multiple views about how a person's wellbeing can be setback. Comparativists claim that the magnitude of harm depends on the difference between V's actual state and some other relevant state with which this is compared (call this the C-state). Comparativism is further divided into two separate categories: temporalism and counterfactualism. Temporalism holds that the relevant C-state is the state V was in prior to the harmful event, and counterfactualism holds that the relevant C-state is the state V might now be in had some other course of events occurred. Temporalism and counterfactualism often take a simple form. A comparative view is simple if it measures the magnitude of harm in any given case with reference to only one C-state. So a simple version of counterfactualism might measure the harm caused to V by some event, E, by comparing her actual state with the state she would now be in had E not occurred. A simple temporalism might measure the harmfulness of E by comparing V's actual state with the state that she was in at t1, immediately prior to the occurrence of E. I will argue that we ought to accept some version of counterfactualism, although a simplistic version such as the above is untenable. To begin with, however, we must address the

arguments that oppose all forms of comparativism. Some of these have recently been pressed by Seana Shiffrin.¹⁴

Shiffrin argues that counterfactual accounts cannot explain the important asymmetry between enduring harm and failing to receive a benefit. Consider two schematic cases: (1) A harms B, reducing B's wellbeing by 5 units and (2) A fails to benefit B by 5 units.¹⁵ In (1), the victim is worse off by 5 units than she would be had she not endured the harm, and in (2) she is worse off by 5 units than she would be had she received the benefit. Shiffrin complains that the counterfactual view treats these cases as identical. More specifically, the worry is presumably that counterfactualism makes these cases look *morally* identical. This threatens the widely held belief that there is an important moral difference between harming and failing to benefit.

This objection is not very convincing. The counterfactual view is not committed to saying that A is harmed in (2) at all. It is committed to explaining all harms in terms of being worse off compared to some C-state; it is not committed to labelling all cases of being worse off compared to some C-state as harms. Whether a proponent of the counterfactual view labels such a case a harm will depend on how the view is further specified. In particular, it will depend on what C-state is chosen as the baseline for comparison with the victim's actual state. Shiffrin objects to one implausible version of the counterfactual view: that which picks as the relevant C-state in (2) the state B would be in had she been benefitted. But why would anyone adopt such an implausible version of the counterfactual view when other ways of

¹⁴ Shiffrin, S. "Wrongful Life, Procreative Responsibility, and the Significance of Harm", at pp. 121 – 3.

¹⁵ Of course, the appeal to schematic units of wellbeing sidesteps some important issues. I am not assuming here that all components of wellbeing are measurable on a common scale, although I believe that some are measurable in this way. Moreover, the measurability of harm depends on the currency. For example, quantitative comparisons are most clearly acceptable for those who believe that harm is setback to resources rather than wellbeing, because many resources such as money and food are straightforwardly measurable.

specifying the baseline are available? For one thing, it ignores the important distinction between harming and failing to benefit! I will return to the baseline issue below.

Shiffrin presents another argument as to why counterfactual accounts cannot explain the asymmetry between harming and failing to benefit. Suppose A has 10 units of wellbeing and B has 14, and this distribution is morally arbitrary. If some event causes A to gain two units and B to lose two, they will end up equally well off. The counterfactual view holds that A is benefitted and B is harmed, but given that they now enjoy the same level of wellbeing, Shiffrin urges, why should we care about the difference?

The objection suggests that we ought not to see A as benefitted and B as harmed because it is a good thing that A and B now have equal wellbeing. The judgement that B is harmed is misleading because this implies that something morally bad has happened, when in fact a more desirable outcome has been reached. This objection confuses *pro tanto* with all-things-considered judgements. Accepting that B is harmed does not entail the conclusion that it is a bad thing that B is harmed. At most, it implies that it is bad in one respect, but this might be overridden by the value of the equal outcome.

There are also many reasons why we might care about the difference between A's benefit and B's harm that make no reference to the fact that A and B now share the same level of wellbeing. The moral status of the outcome may depend on the means by which the distribution is changed. If B is *wrongfully* harmed, she may have a claim for compensation, even though this would not be the case if the distribution was altered by natural causes. It is too narrow to focus only on the fact that A and B now share the same level of wellbeing. This is at best one reason to favour the outcome, which can be defeated by moral considerations relating to how the distribution is altered.

Shiffrin also suggests that the counterfactual view cannot identify harm where it occurs. She believes that if A is made very badly off overall but a small interest of hers is furthered then she is not benefitted: she is just harmed. But if she is made very well off overall, but suffers a moment of intense pain, it *is* true that she is harmed, even though she has enjoyed a net benefit. This thought seems to rest on a confusion between local and net

harms and benefits, a distinction that Shiffrin draws elsewhere. In the above example, the reason it is odd to say that someone is benefitted if they become very badly off overall is simply because they suffer significant net harm. The person who suffers two broken legs, but a slight improvement in her chronic nausea, enjoys a local benefit but suffers a net harm, and this thought seems just as intuitive as the judgement that the person who is made very well off overall but suffers a moment of intense pain suffers a local harm but enjoys a net benefit.

B. Event-based Accounts of Harm

Matthew Hanser has recently presented an alternative view which he calls the event-based theory of harm. This differs from both comparative and non-comparative accounts because it argues that suffering harm is “a matter of undergoing an event whose status as the undergoing of a harm does not derive from the badness of any associated *state* at all.”¹⁶ So whilst comparative and non-comparative accounts require an examination of the victim’s actual state, Hanser’s view shifts the focus onto the event the victim undergoes, independently of the state in which she is subsequently placed. It is fair to say that he offers this view as an alternative to comparativism and non-comparativism. At least, this is implied by his denial of the assumption that these two categories exhaust the options for theories of harm.¹⁷ Hanser emphasises this contrast by dedicating the first part of his article to criticising comparative and non-comparative views, then showing how his own contribution overcomes the problems he associates with them. One thing that will become clear in the analysis of the event-based view is that Hanser’s presentation is misleading. He does not

¹⁶ Matthew Hanser, “The Metaphysics of Harm”, *Philosophy and Phenomenological Research*, Vol. LXXVII No. 2, (2008), 421 – 450 at p. 422.

¹⁷ *Ibid.*

exclude all comparative and non-comparative claims from his theory, making it curiously piecemeal, and rendering it amenable to his own criticisms of counterfactualism.

Hanser begins by adumbrating the kind of event that harm is. The basic type of harm is the loss of some quantity of a basic good (and conversely the basic type of benefit is the acquisition of some quantity of a basic good). Hanser does not give a complete list of these goods but provisionally takes them to include general physical and mental abilities. He then defines other harms and benefits recursively in terms of these 'level-1' harms and benefits. For instance, to suffer a preventive harm is to be prevented from receiving a level-1 benefit (acquiring some quantity of a basic good), and therefore constitutes a level-2 harm. Increasingly complex chains of preventive events can generate nth-level harms and benefits. These conditions establish that someone has been harmed or benefitted and, crucially, are also intended to characterise the badness of harms and the goodness of benefits.

Of course, to establish *that* someone has been harmed is not to determine the magnitude of the harm. One difficulty with Hanser's account is that these two questions implausibly come apart. Suppose A loses 25% of her vision. On this view, she has been harmed because of this loss, and the badness of harm consists in the loss. But, as Hanser is aware, separate considerations must be introduced to calculate the magnitude of the harm. This could depend on both comparative and non-comparative contextual factors. Consider two victims, A and B. A has 25% of her vision, B has 100%, and both A and B lose 25%. A, who is now completely blind, has suffered a greater harm than B, who still retains 75% of her vision, even though they have both undergone the same event. The reason is that A is worse off for losing 25% of her vision than B because she has incurred some ancillary losses that B has not, namely the complete inability to detect light, movement and so on. These losses are relevant to measuring the magnitude of the harm A has suffered and they are discovered by examining A's current *state*, either in an absolute sense or by comparison to a relevant alternative.

Hanser recognises this, and offers a series of additional considerations that determine the magnitude of a harm. These include the duration of its bad effects; how much

of a basic good is lost and how much is retained; and whether its bad effects are non-comparatively bad. In the aforementioned case, the fact that B retains 75% of her vision while A is completely blind explains why A suffers a greater harm than B. But the addition of these considerations, though clearly necessary to achieve plausible results, seems *ad hoc*. Unlike the conditions that establish *whether* someone has been harmed, which appeal only to events, these considerations appeal to both states and events. For example, some harm-events will put the victim in a bad state, and the magnitude of the harm depends on the duration of this state. So in fact Hanser's account is only event-based with respect to whether harm occurs, not with respect to its magnitude. With respect to the latter, his view takes into account a mixture of state and event-based factors.

Furthermore, recall that the conditions that determine whether someone has been harmed are also meant to capture the badness of harm. Suppose that A and B both lose the use of their legs and thereby lose the same quantity of a basic good. A is in this state for a day while B is in this state for a year (although they both ultimately receive treatment that restores the use of their legs). On Hanser's view, the fact that A and B have been harmed is established by the event of losing the use of their legs. This also captures the badness of what has happened to them. It is important for Hanser that these answers are generated only by event-based factors, as he wishes to avoid problems he associates with state-based determinants, most notably problems posed by death. The *magnitude* of the harm, however, is determined partly by the duration of the state in which this event causes them to be (B is in a bad state for longer than A, so he suffers a greater harm) – a state-based rather than event-based factor. But given that only event-based factors are meant to characterise harm, the account implies that the factors that determine the magnitude of harm have nothing to do with the badness or nature of harm.¹⁸

¹⁸ Actually, as I mentioned, on Hanser's account the magnitude of harm is determined by both event-based and state-based factors. But this complication is irrelevant to the criticism: the

Judith Thompson raises a similar objection in her reply to Hanser's article. According to Thompson, Hanser's account implies that:

'The badness of an event that consists in losing the use of one's legs is independent of, does not rest on, is not due to, the badness of its end-state, namely the state that consists in lacking the use of one's legs. How can Hanser consistently say that and say also that the badness of the event that consists in a man's losing the use of his legs is the greater if it issues in his being, for a longer time, in the state of lacking the use of his legs?'¹⁹

In reply to Thomson, Hanser states:

'I distinguish sharply between (a) that in virtue of which something counts as a harm, and (b) that which determines the seriousness or severity of a harm. I argue that bad states do not enter into a proper account of the former... But when a harm is nonlethal, and so does leave its subject in a bad state, I grant that the duration of this state can affect the seriousness or severity of the harm the subject undergoes in being put into that state. I see no outright inconsistency here.'²⁰

It is true that Hanser's view involves no outright *contradiction*. This is because Thomson is probably wrong to impute to Hanser the claim that the duration of an end-state affects the badness or nature of harm (although Hanser's text is a little unclear on the issue).

unwanted implication is that *some* factors (the state-based ones), which help determine the magnitude of harm, have nothing to do with the badness of harm.

¹⁹ Judith Jarvis Thomson, "More on the Metaphysics of Harm", *Philosophy and Phenomenological Research* Vol. LXXXII No. 2, (2011), 436 – 458, at p. 457.

²⁰ Matthew Hanser, "Still more on the Metaphysics of Harm", *Philosophy and Phenomenological Research* Vol. LXXXII, No. 2, (2011), 459 – 469, at p. 467.

Hanser would be contradicting himself if he made this claim, as he previously held that the badness of harm is wholly captured by undergoing a certain event, independently of state-based factors. But Hanser does not claim that the duration of the end-state affects its badness, he claims that it affects its seriousness or severity, thus appealing to a distinction between the badness and its severity. But although his view is not contradictory, it *is* arbitrary. For how can the severity of harm have nothing to do with its badness? This would mean that two harms can be equally bad, even though one is much greater than the other! It is just implausible to characterise the badness of harm independently of its magnitude.

Furthermore, recall Hanser's definition of harm as the loss of some quantity of a basic good. As mentioned, these losses may vary in their severity. I suffer a greater harm by losing 75% of my vision than 25%. Do *these* differences in severity affect the badness of harm? Hanser *could* answer in the affirmative without contradiction because these considerations are event-based, determined by the nature of the event rather than the resultant state. But event-based or not, they are still considerations of *severity* or *magnitude*. So it would be arbitrary to allow the badness of harm to be characterised by event-based considerations of severity but not state-based ones. If, on the other hand, Hanser were to answer in the negative and hold that they do not affect the badness of harm, then the gulf between the badness harm and its magnitude is widened even further.

Hanser is motivated to develop his event-based view largely because he believes that state-based views have serious difficulty dealing with death.²¹ It is therefore worth asking whether his view fares any better with respect to death. We can ask: what, on Hanser's view, is the magnitude of harm caused by death? He accepts both that the quantity of basic goods that is lost and the bad effects of a harmful event determine the magnitude of a harm, but how do these factors interact? Death involves the total loss of goods, but has no further bad effects, because as Hanser argues being dead is not bad *for* the deceased – the state of being

²¹ See Matthew Hanser, "The Metaphysics of Harm", pp. 437 – 40 and "Still More on the Metaphysics of Harm", pp. 463 – 4.

dead has no wellbeing value at all for the deceased. On the other hand, being blinded in one eye for the rest of one's life only involves the partial loss of one good, but the bad effects of this event have a much greater duration. So what explains why death is more severely harmful than partial blindness? The fact that the magnitude of harm can be affected by both event-based and state-based factors opens up a new realm of complications for Hanser, as he must now provide an explanation of how these factors interact which plausibly orders harms according to their magnitude.

I have already noted Hanser's non-exhaustive list of factors that help determine the magnitude of harm. This list does not include reference to counterfactual states (although, as I mentioned, given the fragmentary nature of his view, whereby events exhaustively determine when harm occurs and capture its badness, and both events and resultant states determine the harm's magnitude, Hanser does not commit himself to rejecting the relevance of counterfactuals here). But counterfactual states do seem to be relevant in some cases. If, had V not been killed by E, she would have been killed by E2, V's death harms her less overall than if she would have lived a long life if not for E. If he accepts that these counterfactuals are relevant, the event-based view is to some extent reliant on counterfactualism, so it is not clear how the former fares better with death than the latter.

Hanser's claim that the duration of a bad state affects the severity of harm, in the absence of a detailed explanation of how the various factors go in to measuring this severity, introduces other dilemmas into the theory. To demonstrate, suppose A loses all of her basic goods and dies, and B loses all of her basic goods and falls into a permanent coma. What does Hanser's view imply in this case? A plausible conclusion is that A and B are harmed to the same degree. Admittedly, B is placed in a bad state for a much longer duration (since she is still alive, her state of having no basic goods persists) but this does not make it true that she is harmed more than A. Hanser could reach the right conclusion if he maintained that the severity of the harm is the same because the two individuals lose an equal quantity of basic goods. But we have already noticed that a medley of factors, including the duration of the bad state, go into determining severity. Here we have a case in which the duration of a bad

state does *not* affect the magnitude of harm, so Hanser owes an explanation as to how this is distinct from other cases where the duration of the resultant state *does* affect severity. Alternatively, if Hanser replies that, since B is still alive, the duration of her comatose state increases the severity of her harm, the view implies that B's harm is much greater than A's. This is the wrong result: falling into a coma is not *more* harmful than death, although it is arguable that they are on a par.

Finally, some attention must be paid to Hanser's refinement of his event-based theory in his reply to Thomson. 'Whenever one of a subject's states impairs his functioning' he says, 'there is arguably an ongoing event of that state's impairing his functioning.'²² In other words, all harmed states are describable in terms of events, so instead of saying that A has been harmed because she has been placed in the state of having a broken leg, we can say that she has been harmed because of the ongoing event of her broken leg preventing her from walking. To some extent, this unifies Hanser's theory, as now he can determine both *that* someone has been harmed and the severity of the harm solely with reference to event-based factors. One problem with this move is that it relaxes the definition of 'event' to the extent that it drastically expands the class of things that can *cause* harm. It is plausible that one can be harmed by a person or a falling rock, but is it plausible that one can be harmed by the internal state of having a broken leg? Isn't this an unnecessary intermediary description? Suppose I break your leg. We now have two different descriptions of the situation: (1) my action puts you in a state of impaired functioning and the ongoing event of this state's impairing your functioning is a harm; (2) My action puts you in a state of impaired functioning and this state is a harm. It seems clear that (2) is a simpler and less convoluted description than (1).

It is also hard to square this refinement with his definition of harm as the loss of a quantity of a basic good. Hanser's amended theory implies that throughout the period when A has a broken leg, he is in the process of being harmed. But intuitively, the loss of a basic

²² Matthew Hanser, "Still More on the Metaphysics of Harm", p. 468.

good can be distinguished from its bad effects – my losing my leg is not the same as the impairment caused by this loss. In fact, in his original article, Hanser makes this distinction himself in separating the undergoing of a harm from its bad effects. Hanser’s expansion of the definition of ‘event’ now evaporates this distinction.

There is a final problem with the definition that Hanser’s refinement could have avoided, but unfortunately does not. It is odd that harm should be defined as the loss of a *basic* good. Hanser does not give a summary of basic goods, but in describing them as basic he distinguishes them from non-basic goods, the loss of which does not constitute harm. But why should the loss of a non-basic good not constitute harm? Hanser’s event-based account is susceptible to a similar criticism he himself makes of non-comparativism. Non-comparativism wrongly implies that those who suffer a decrease in wellbeing but are still non-comparatively well off have not been harmed. To see the plausibility of this argument, consider a Nobel Prize winner who suffers a modest decrease in her intellectual abilities as a result of a stroke. She is less intellectually capable than she was, but she is still more intelligent than the average person. Even though she is not badly off in non-comparative terms, her stroke still harms her. On Hanser’s view, one can suffer the loss of a non-basic good, with a corresponding decrease in wellbeing, without being harmed. Possessing exceptional rather than above average intellectual abilities is a non-basic good if anything is, so Hanser’s view also implies that the Nobel Prize winner who suffers a stroke is not harmed.

What are we to make of Hanser’s view in light of this analysis? It comes across as arbitrary and piecemeal in its insistence on the distinction between the identification and badness of harm on the one hand, which are captured solely by the event of losing some quantity of a basic good, and the severity of harm on the other, which is calculated according to a mixture of event-based and state-based factors. This structure undermines one of the primary rationales for developing the event-based view – the difficulty of competing theories to assess death as a serious harm – by shifting the problem to a different level of the theory. While counterfactual views have difficulty holding that death is a harm at all, the

event-based view faces an equivalent difficulty assessing the severity of harm caused by death. These shared difficulties have a common source: the role of counterfactual states in determining the measure of harm. Since Hanser does not preclude the possibility of counterfactual states playing this role (and they do play this role in some cases) it is not clear that his view fares any better with respect to death.

Perhaps Hanser's view can be developed. His attempt to do so in subsequent work brings the theory closer to state-based views. But as it stands, it has many moving parts. It is not exclusively event-based, for the reasons given above, and it has the potential to include both comparative and non-comparative claims. Despite his presentation of it as an alternative to the current set of options, therefore, Hanser's theory may turn out to make four separate types of claims: event-based, state-based, comparative and non-comparative! As it is doubtful that it really deals better with death than its competitors, we have reasons of parsimony to reject it. Of course, Hanser is right that all else is not equal as there are other problems with non-comparative and simple comparative views. But as I will now argue, these can be overcome by developing a complex version of comparativism.

C. The Complex Counterfactual View

What are the problems afflicting other views?²³ On non-comparative accounts of harm, setback to wellbeing means placing someone in a non-comparatively bad state, rather than a state that is worse than some relevant alternative. We have already encountered a central problem with non-comparativism: it suggests that there are cases where an event, E, renders V worse off, but because V ends up in a relatively good state, she is not harmed. This implication is implausible in cases where V is placed in a non-comparatively good or neutral

²³ For discussion of these and other objections to non-comparativism, see Victor Tadros, *What Might Have Been*.

state by being harmed (consider again the Nobel prize winner who suffers a modest decrease in her talents as a result of a stroke).

There is a related objection: someone can also be made better off and still be badly off. That is, it is possible to put a person in a non-comparatively bad state by *improving* her condition. This possibility yields an even more counterintuitive implication for non-comparativism. Suppose B has a severe headache and a doctor gives her a painkilling drug, causing her to be in the non-comparatively bad state of having a moderate headache. The doctor has not harmed B, she has benefitted her, but it is not clear that non-comparativism can explain why. After all, the doctor's act causes her to be in a bad state. The answer, of course, is that B is still better than the relevant alternative. But if harm is determined solely by examining B's actual state, without comparing it to some other state, how can it be relevant that the doctor improves her condition relative to her previous state or to that B would be in had she not intervened?

Comparative accounts yield the correct implications in these cases. A has been harmed because she is worse off than she would otherwise be, or worse off than she was before she was given the drug; B has been benefitted because she is better off than she would otherwise be, or better off than she was before the doctor intervened. However, simple counterfactual theories face difficulty in dealing with cases of preventive harm, where causing harm to a victim prevents equal or greater harm being imposed on her. For example, if A breaks B's arm, but had she not broken B's arm, C would have broken B's arm, the simple counterfactual view implies that A does not harm B by breaking her arm. This is a counterintuitive result. We want our theory to pick out certain harms even when they pre-empt equal or greater harms.

A complex counterfactual theory allows that different C-states are relevant for different judgements. In the above case, such a view might hold that A harms B by breaking her arm because B is worse off now than she would have been had no one broken her arm. A challenge for the complex counterfactual view is to formulate non-arbitrary principles to establish the appropriate C-states for different judgements.

Judith Thomson rejects non-comparative accounts for similar reasons to those sketched above, and in response to the criticisms of simple counterfactualism she develops a complex view that she calls the Revised Counterfactual Comparative Account of Harming.²⁴

According to this account:

A harms B just in case B is in a state *s* such that:

A causes B to be in *s*, and for some state *s**

(a) A prevents B from being in *s** by the same means by which A causes B to be in *s*, and

(b) B is worse off in a way for being in *s* than he would have been if he had been in *s**

Thomson claims that this view has plausible implications in relevant cases. In one example, she supposes that Villain C throws acid in A's eyes to blind him. The acid begins to affect A's eyes, but before it can complete the process of blinding A, bystander B intervenes, using a neutraliser on A's eyes, causing A to be in a state short of blindness – that of having dim vision. In this case, although both C and B cause A to be in 'has dim vision', only C harms A. The explanation, Thomson thinks, is that C causes A to be in 'has dim vision' by throwing acid in his eyes, and this same event also prevents A from being in 'has good vision'. The bystander's intervention, on the other hand, does not prevent A from being in 'has good vision', or indeed any other state in which A would be worse off. Therefore, we can conclude that the villain harms A but the benevolent bystander does not. Thomson's view has the same result in equivalent cases such as the severe headache case mentioned above.

This view has the following implication: A does *not* harm B if A causes B to be in *s*, and for some state *s**, A prevents B from being in *s** by the same means by which A causes

²⁴ Judith Jarvis Thomson, "More on the Metaphysics of Harm", pp. 12 – 13.

B to be in s , and B is *not* worse off for being in s than she would have been if he had been in s^* . But this means that her view does not solve the problem of preventive harming. It implies that, if A breaks B's left arm, preventing C from breaking B's right arm, A does not harm B because A's act (breaking B's left arm) both causes B to be in s (having a broken left arm) and prevents B from being in s^* (having a broken right arm), and B is not worse off for being in s than she would have been if she had been in s^* . Consider another possible state in this case: B having no broken arms. Call this state s^{**} . It is also true of s^{**} that A prevents B from being in s^{**} by the same means by which A causes B to be in s , and B is worse off in a way for being in s than he would have been if he had been in s^{**} . Therefore, on Thomson's view, A both harms B and does not harm B.

Moreover, this view yields false positives. Suppose A gives B £10, thus preventing her from receiving £20 from a charitable egalitarian, who then decides to give her money to someone who has not recently been benefitted. A's act places B in one state (having £10), but simultaneously prevents her from being in another (having £20) in which she would be better off. B is worse off in her actual state than she would be in this other state that A prevents by placing her in her actual state, and therefore A harms B. This is an unwanted result. The intuitive judgment is that A benefits B by giving her £10.

The problem is that, in order to decide whether a person has been harmed in a given case, Thomson's view allows us to pick any C-state as long as it is prevented by the putative harmful act. But there are many counterfactuals that are prevented in this way and the victim's wellbeing may vary significantly between them, so the harm will be indeterminate. Another way to look at this problem is that in some cases Thomson's view does not identify the relevant baseline. In the charitable egalitarian case above, the reason that A benefits rather than harms B by giving her £10, even though this prevents her from receiving a greater amount, is that the appropriate baseline for comparison is the amount of money B has without any charitable acts. Compared to this baseline, A benefits B by making her £10 better off.

Now compare this to another case. Suppose A does not like B and wishes to minimise B's wealth. A knows that C, the charitable egalitarian, will give B £20 but only if B has not recently been given a sum of money. A gives B £10 in order to prevent C giving her £20. In this case, the same preventive act of giving B £10 seems like a harm rather than a benefit. The baseline for comparison is not B's holdings prior to any gifts, but the money she would have if C had donated to her. This suggests that the intention of the agent affects our choice of the baseline for comparison. If this is true, it follows that intentions can affect whether a person is harmed or benefitted by the same act. The principles that govern what counterfactuals to take into account are not adequately captured by Thomson's view. No guidance at all is given as to how we should pick out the relevant C-states.

There is another problem, which we can call the problem of scope. The possible consequences of a given act are potentially limitless. Some consequences may be realised long after the act has taken place. Other consequences, although not temporally distant, may be wildly unexpected. In English law, the test for 'remoteness' limits the scope of the damage for which a tortfeasor can be liable. The test is that of 'reasonable foreseeability', and so it is not directly tied to temporal considerations, although these may be relevant. There are two components to this challenge, which we can call the *relevance problem* and the *epistemic problem*. The relevance problem is summed up in the following question: on the counterfactual theory, which C-states are relevant for comparison with the victim's actual state? This way of stating it clarifies that the baseline issue is only one component of a larger problem. It is not only necessary to pick out the relevant baseline for comparison, it is also necessary to specify the requisite scope.

For example, in standard cases where A causes B some physical injury, what is the temporal location of the relevant C-state? The duration of the victim's suffering may be uncertain. Perhaps the scope should be limited to the point at which the physical injury is removed or disappears, assuming this is possible. Even so, the initial injury may have other effects that persist after the physical injury is gone. Suppose B's injury prevents her emigrating from England to Australia. Apart from the injury itself, should we take into

account the difference in B's wellbeing between her life in England and the life she would lead in Australia in calculating the magnitude of harm A causes B?

This leads to the *epistemic problem*. The larger the temporal scope for deciding questions of harm, the more difficult it will be to ascertain the relevant counterfactual information. It may be impossible to make any satisfactory judgments about a person's wellbeing in the distant future. In the above case, we simply have insufficient information about the quality of B's life in Australia compared to her life in England, even though there may be a fact of the matter as to which life is more conducive to her wellbeing.

Asking whether an agent has been harmed in a given case is therefore underspecified. An agent's act may be a causal condition for a wide variety of consequences, which vary in terms of their effect on the victim's wellbeing. If, in the above example, the victim would have been killed by a bolt of lightning had she moved to Australia, taking into account this consequence would generate the conclusion that she is benefitted. If she would have met her future partner at the airport, perhaps she was harmed by the act that prevented her from going. But in both cases it is plausible that we should ignore these counterfactuals when calculating harm because they are mere coincidences. A similar problem arises in the philosophy of causation, as A's act may be a factual condition of B's harm without causing it in any normatively significant sense.²⁵ In the context of causation, we need to distinguish between relevant conditions and mere coincidences when attributing causal relations. In the present context we need to do the same to determine which C-state should be considered when assessing harm.

Recall that Thomson's view, when applied to cases of preventive harm, has the implication that A both harms and does not harm B. As it stands, this is an unacceptable conclusion, but it contains a kernel of truth. We should extend the distinction between *local* and *net* harm, and say that A causes B local but not net harm. A causes B local harm if A places B in a state in which B is worse off than the state in which nothing makes B worse

²⁵ See Victor Tadros, *Criminal Responsibility*, Oxford: Oxford University Press, (2005), pp. 179 – 81.

off. This captures the very general set of cases in which A has some local negative impact on B's wellbeing. So, for example, if A breaks B's arm in order to prevent C breaking both of B's arms, A causes B local harm because B is worse off with a broken arm than she would be if neither A nor B (nor anything else) made her worse off. We can then compare the local harm A causes B with the net harm B incurs, introducing other relevant counterfactuals.

The distinction is important for a number of reasons. First, it is clear that in some cases it is permissible to inflict harm on a person in order to prevent greater harm being inflicted on her. The permissibility of A harming B to prevent greater harm being inflicted by C depends on at least two factors: the level of harm A will inflict on B and the level of harm C will inflict on B if A does nothing.²⁶ For example, if A can only pre-empt C by inflicting on B a greater degree of harm that C would inflict, it is not permissible for A to harm B. The relevance of these questions about the magnitude of harm suggest that any adequate theory must be able to say that A harms B in this case and also determine the magnitude of that harm. These distinctions, of course, can only be drawn by the complex counterfactual account.

More generally, as Shiffrin notes, the fact that local harm has been inflicted on a person, even if that person has not suffered net harm, may give us reasons that would be obscured by the failure to distinguish between local and net harm. If A saves a person from drowning at the expense of breaking her arm, the mere assertion that A has benefitted her may obscure the fact that A now has reason to set her broken arm.²⁷ This reason arises naturally from the thought that she has been harmed by having her arm broken, and this thought is not available without the above distinction, as she is better off with the broken arm than she would be had she not been saved.

²⁶ On this point, see also Victor Tadros, "What Might Have Been".

²⁷ See Seana Shiffrin, "Wrongful Life, Procreative Responsibility and the Significance of Harm", pp. 120 – 1.

This discussion shows that the metaphysics of harm, on the complex view, depends on some unusual factors. I cannot provide an exhaustive list here, but one relevant factor is the mental state of the injurer. The intentions of an agent who inflicts an injury seem particularly important. If A's intention in breaking B's arm is to prevent C breaking both of B's arms, then it is appropriate to consider the counterfactual possibilities that fall within the scope of A's intention. We should take into account what C would have done had A not intervened, intending to avert this consequence. The result will be that although A causes B local harm, she benefits her overall. The relevant baseline for overall harm is the state B would be in had A done nothing. If, on the other hand, A's intention in breaking B's arm was not to interfere with C's plan, but rather to make B emigrate, so that she could no longer apply for a job that A was pursuing, then perhaps it is appropriate to broaden the scope to include the way B's life will go after she emigrates. In this case, the appropriate baseline is the life B would have enjoyed had she remained in the country, so A causes net harm to B.

Other mental states besides intention are relevant. Suppose that, although A breaks B's arm with the intention to prevent greater harm, A acts with the foresight that B will emigrate and her life will go worse overall. Or suppose instead that, although A intends to prevent greater harm to B and does *not* know that she will cause B to emigrate, she reasonably *ought* to have this knowledge and so is negligent in failing to take this into account. In both of these cases, perhaps we ought to say that A causes B net harm, even though A prevents greater harm being inflicted by C, because of the presence of foresight or negligence as to the long term effects of A's intervention. These factors lead us to broaden the scope of the calculation, yielding the result that A causes net harm to B. Conversely, the fact that harm to V is an unforeseeable result of E does not mean that E does not harm V. If a doctor gives a patient a course of medication designed to improve her condition that unforeseeably makes the patient worse, it is still true that the drug has harmed the patient, even if we would not hold the hospital liable to pay compensation.

So on the view sketched above, there is no single right answer to the question: did A harm B? The question is simply underspecified. It might be objected that there *is* a single

answer. If we aggregate B's wellbeing over the course of his life after A's act, then A harms B if and only if B's aggregate wellbeing is lower than it would be had A not acted. This, it might be urged, is the single factual answer because death is the point at which a prior act can no longer affect one's wellbeing. This assumes that posthumous harms are impossible, which is contestable, but even if this is the right way to understand the relationship between harm and death, it is an arbitrary move to make. When calculating harm, we are rarely interested in such remote consequences, even if we are able to determine them. It is true that the factual question, 'how much harm has A suffered?' is distinct from the normative question, 'is someone liable to compensate A for this harm?' This follows from the fact that we can admit that A has been harmed without thinking it appropriate to impose liability on anyone. Nevertheless, the normative conclusions we *might* draw on the basis of the factual question delimit the scope of the factual analysis. For example, if we might conclude that the harm is wrongful or that the person who caused the harm should compensate the victim, then it is appropriate to restrict the scope of the factual question to consequences the perpetrator intended, could foresee, or reasonably ought to have foreseen.

Finally, to continue the parallel between harm and causation, it is worth noting that the complex counterfactual theory gains support from contrastive accounts of causation, such as those developed by Jonathan Schaffer and Christopher Hitchcock.²⁸ On Schaffer's version of this view, causation involves four causal *relata* rather than the traditional two of cause and effect. Instead of claiming that c caused e, we ought to claim that c rather than some alternative cause, c*, caused e rather than some alternative effect, e*. The complementarity between this view and the complex counterfactual theory of harm is that

²⁸ See, for example, Christopher Hitchcock, "The Role of Contrast in Causal and Explanatory Claims", *Synthese*, 107, (1996), 395 – 419 and Jonathan Schaffer, "Contrastive Causation", *Philosophical Review*, 114, (2005), 327 – 358.

both allow us to pick the relevant counterfactuals to establish causation or harm depending on the context.

3. Conclusion

I have defended wellbeing as the correct currency of harm against competing candidates such as preferences and interests, and endorsed a complex version of the counterfactual view over both non-comparativism and Matthew Hanser's event-based account. We can summarise the positive thesis of the chapter as follows: an event, E, harms V if and only if E reduces V's wellbeing compared to a relevant counterfactual state, S, that V could be in.

Two things should be noted about this view. First, it is incomplete. It remains an open question which counterfactuals are relevant for assessing harm. I have suggested that mental states such as intentions are relevant and coincidences are not, but I have not explored either of these possibilities in much detail or investigated other salient C-states. Second, the view allows that there can be more than one relevant C-state in a given case, thus incorporating the distinction between local and net harm. This enables it to capture our intuitive but apparently contradictory judgements in some cases of preventive harm. Where A harms B in order to prevent B suffering a greater harm at the hands of C, the complex counterfactual view rightly holds that A causes B local harm but benefits her overall.

This position, although in need of greater development, seems to do the best job of tracking the subset of permissibility-affecting facts that underlie our judgments about harm, which was the aim set at the beginning of this inquiry. Identifying the currency as wellbeing distinguishes harms from other subsets of permissibility-affecting facts. This is an important task as it clarifies the different reasons that go into making acts permissible or impermissible. It seems there is a difference, for example, between reducing a person's

wellbeing and interfering with them.²⁹ Paternalistic acts are not always harmful as they can improve a person's wellbeing, but they do constitute interference. If some of these acts are impermissible, therefore, this is not because they are harmful.

The complex counterfactual component is also an improvement over its competitors. Non-comparativism is insensitive to the way in which a person is placed in a non-comparatively good or bad state. A can be put into a non-comparatively *good* state by being made *worse off* (the literary genius who has a mild stroke) or a non-comparatively *bad* state by being made *better off* (the person in severe pain who takes mild painkillers). Clearly it is impermissible to induce a stroke in the literary genius and permissible to give the patient painkillers. The best way to explain this is that the novelist is harmed and the patient is benefitted. Non-comparativism cannot explain these judgements in this way because it implies the converse: that the novelist has been benefitted and the patient has been harmed. Comparativism can explain these judgments in the right way.

Similarly, the complex view is an improvement over simple counterfactualism. The latter implies that if A injures B, preventing B being struck by a bolt of lightning, A does not harm B because B would have been even worse off had A not acted. But it is not permissible for A to injure B, assuming he is not doing so in order to prevent greater injury, and this is hard to explain without the judgment that A harms B. Only the complex view can imply that B is harmed in this case, allowing us to explain the fact that A's act is impermissible.

Finally, complex counterfactualism should also be preferred to Hanser's view. In my critique I noted some internal problems with the event-based theory: it is arbitrary, unnecessarily complex and has some implausible implications. Importantly, it is also incomplete in the sense that Hanser does not specify whether counterfactuals are relevant in calculating the magnitude of harm. If they are, then Hanser's view suffers from the same problems as counterfactualism generally. If not, then it does not capture important permissibility-affecting facts about harm. Suppose that a child and a 100-year-old patient in

²⁹ I will explore this difference in greater detail in the context of corrective action. See chapter 4.

the final stages of terminal cancer catch an infection that will kill them if it is not treated. A only has time to treat one of the patients. Most people will say that A ought to save the child. Perhaps, more strongly, it is impermissible for A to save the 100-year-old rather than the child. This judgement clearly depends on the idea that the child will be harmed by the death more than the older patient, because the child has many more years to live and the older patient would have died immanently anyway. If Hanser's view cannot take into account these counterfactuals, it fails to explain these judgements about the doctor's actions.

Corrective Action

In the previous chapter we developed a theory of harm. This held that harm consists in setback to wellbeing and ought to be measured in terms of the difference in wellbeing between the victim's actual state and a counterfactual state the victim might now be in. This theory has general application, but from the perspective of normative tort theory its key benefit is that it allows us to identify the kind of consequences that typically merit corrective action, and make judgements about what actions are appropriate given the degree of harm done. For example, the most familiar form of corrective action – compensation – is usually given in response to harm and usually quantified according to the magnitude of harm.

Most theoretical discussions about liability in tort law focus on duties to compensate, or more specifically duties to pay monetary damages. In principle the scope of corrective duties is broader than this, however.¹ One danger of focussing solely on duties to compensate, or simply referring to corrective or remedial duties without further analysis, is that this may obscure a more fundamental distinction between different categories of corrective action.

Some theorists have distinguished between different types of corrective action. Ernest Weinrib holds that private law *primarily* corrects normative rather than factual gains and losses, which consist in deviations from regulatory norms rather than setbacks to welfare. Robert Goodin distinguishes between means-replacing and ends-displacing forms of compensation, drawing on the important difference between giving a victim equivalent means to pursue her ends and giving her new ends which positively contribute to her wellbeing. There is, however, one distinction that theorists of private law have not yet

¹ Of course, it is also broader than this in practice, a fact that the focus on compensatory duties tends to obscure. For an argument that corrective justice cannot explain the various remedies available to tort claimants, see Benjamin Zipursky, "Civil Recourse, Not Corrective Justice", *Georgetown Law Journal*, 91, (2003), 695 – 757, at pp. 710 – 13.

explored: between *negating* and *counterbalancing* detrimental states. I define these terms in the following way (X refers to the currency of the corrective duty; that is, the good or goods that have been setback by the detrimental state): a negatory duty renders the victim's future X *identical* to what it would otherwise be, and a counterbalancing duty renders the victim's future X *equal* to what it would otherwise be. I argue that these should be recognised as two fundamental categories of corrective action, for two main reasons. First, the failure to do so causes confusion in theoretical discussions of corrective duties; and second, there is a morally significant difference between the two types of corrective action.

I understand corrective duties broadly to involve any action that is corrective in nature, such as removing harm, paying damages, apologising and so on. Similarly, the detrimental states to which corrective measures are directed may be wide-ranging, including harms, non-harmful loss and the threat of future harm.² Three initial points should be made about this description. First, it is intended to be *freestanding*. It is neutral about (1) what duties are, (2) how corrective duties are related to primary rights and (3) what corrective duties exist. The description is not dependent on any *analysis* of duties and should be consistent with different semantic views about what duties are. Furthermore, it is committed neither to the view that all corrective duties are correlated with primary rights, nor to the view that some are owed independently of rights. It is also consistent with a wide range of normative views about what corrective duties exist and how they are generated. In particular, it does not assume that corrective duties only arise from *wrongs*. Corrective duties that arise from wrongs are usually modelled with a binary distinction between primary and secondary

² For now I will assume that removing a threat of harm is a corrective action. This assumption is false if corrective actions only respond to decreases in wellbeing, since threats in themselves are not harmful (unless of course the threat is communicated, in which case it can induce fear and panic). Duties to remove threats, however, are just as important as duties to compensate for harm *ex post*, and for this reason corrective duties may not be a category of fundamental interest. I discuss this issue further in chapter 7.

duties.³ A primary duty is a duty to avoid committing a wrong. A secondary duty prescribes a corrective measure in response to some consequence of the breach of a primary duty. Secondary duties are, of course, an important category of corrective duties, but at the conceptual level we must be open to the possibility that the latter arise *strictly*, independently of wrongdoing.

Second, the duties I am discussing are moral rather than legal. Although corrective duties figure importantly in private law, and I will occasionally refer to legal practices when discussing them, they occupy a broader territory. One reason for this is that some corrective duties may not be sufficiently important to justify the cost required to legally enforce them. Another reason is that we may owe corrective duties to things that cannot be party to a private action, such as non-human animals. If it is a serious moral wrong to torture a sentient being for pleasure, we may incur a corrective duty by committing this wrong, even if the thing we torture is a non-human animal. Although I will not discuss this possibility any further, it cannot be ruled out by conceptual stipulation. *If* a duty to aid a non-human animal that one has injured exists, it is a corrective duty.

Third, as I have already indicated, a corrective duty has two components, which I call the *action* and the *currency*. The action is the corrective measure prescribed by the duty.

³ It may be possible to extend primary duties to cover cases where a defendant is strictly liable. The primary duty would then contain an outcome requirement. In other words, it would be a duty not to cause harm, or some other detrimental state, rather than a duty not to commit a wrong, irrespective of whether the wrong causes a negative outcome (although note that some theorists think of the duty of care in negligence as having both a conduct *and* an outcome requirement, for example Arthur Ripstein and Benjamin Zipursky, "Corrective Justice in an Age of Mass Torts" in *Philosophy and the Law of Torts*, Ed., Gerald Postema, Cambridge: Cambridge University Press, (2001), pp 214 – 49 at p. 220.) I leave this possibility aside here, largely because it may be more plausible to think of duties in strict liability as duties not to cause some outcome *without correcting it*, rather than duties not to cause an outcome *per se*.

If a corrective duty requires the duty-bearer to pay £X to a victim, the payment of £X is the action. Actions may be described specifically or generally. The payment of £X is a specific description, but the payment of monetary damages in law is a general description of a type of corrective action. I will not discuss monetary compensation directly in this paper because, as I will argue in the penultimate section, payment of damages often serves both a negatory and a counterbalancing purpose.

The currency is the good or goods that have been setback and to which the corrective action responds. If a corrective action responds to the victim's decrease in wellbeing as a result of physical pain, then *wellbeing* is the currency. There may be disagreement about the appropriate currency or currencies of corrective duties. To simplify, I will generally consider cases in which wellbeing is the appropriate currency, but this is not the only available view. If D steals £10 from V, a millionaire, the loss may have no effect on V's wellbeing. We may still think that D owes a corrective duty, in which case *resources* rather than wellbeing are the relevant currency. Individual cases may also involve setbacks to multiple currencies, such as resources *and* wellbeing. This would be true if D injures the millionaire as well as stealing her money. Another more general view, endorsed by Robert Nozick, is based on the *hypothetical preferences* of the victim. On this view, whether D has compensated for V's harm depends on whether V is given some benefit that renders her indifferent between (1) suffering the harm and receiving the benefit and (2) not suffering the harm.⁴ I will not adjudicate between these possibilities here, since my main focus is on the actions rather than the currencies of corrective duties. This omission is not problematic since the arguments I make about negating and counterbalancing can be motivated irrespective of which currency we adopt.⁵

⁴ See Robert Nozick, *Anarchy, State, and Utopia*, pp. 57 – 8.

⁵ In the previous chapter I argued that wellbeing is the currency of harm. This indicates that it is usually the currency of corrective duties too. Although my examples involve wellbeing, here I remain

Here is how the chapter proceeds. In section 1 I outline some previous conceptual accounts of corrective measures. Robert Goodin's distinction between means-replacing and ends-displacing compensation comes closest to that between negating and counterbalancing defended here, although some modifications are required. He is correct to place emphasis on the victim's ability to pursue her ends, but the effect of an event on the victim's end-setting capacity is not the only relevant factor for understanding the basic categories of corrective action. In section 2 I describe the preferred distinction in more detail, rendering precise the difference between being made identically off with respect to some currency and being made equally well off. Sections 3 – 5 adduce three arguments to support the importance of the distinction: first, the distinction clarifies the debate about whether imposing secondary duties on those who breach primary duties is justified; second, it permits us to describe a more nuanced set of possibilities as to what corrective duties *might* exist, so failing to recognise it obscures important normative questions; and third, we have basic moral reasons to care about the difference between being identically off and as well off. The first two arguments establish that the distinction helps to frame normative questions and avoids confusion, and the third seeks to demonstrate its moral significance. In the penultimate section, I then outline some ways in which the discussion helps us explain and justify specific legal remedies, and I conclude by noting some exceptions to the priority of negating over counterbalancing.

1. Previous Accounts of Corrective Measures

Let us begin by surveying two accounts of corrective measures – those of Ernest Weinrib and Robert Goodin – to pave the way for a new one. As I explained in chapter 2, Weinrib

neutral about the success of the arguments in chapter 3 as my current claims do not depend on them.

invokes an Aristotelian conception of corrective justice according to which the grounds of liability are understood in terms of the annulment of wrongful gains and losses. These gains and losses are essentially normative and consist in the violation of rights. Private law rectifies them by affording the claimant a legal action.

Weinrib seems implicitly to think of compensation as a response to *factual* losses. He cannot think that we compensate only for normative losses, because when he dismisses compensation as a justification for private law, he relies on the idea that we could compensate victims regardless of whether their injuries were wrongfully caused,⁶ and given that there are no normative losses where there are no wrongs, non-normative losses must be amenable to compensation. But compensation can be *involved* in the correction of a normative loss. When a wrong is committed, the law rectifies this by “an award of damages that quantifies the wrong by quantifying the value of the welfare of which the plaintiff has been deprived”.⁷ In this case, correcting the normative loss entails the compensation of factual losses, but is not identical to it. But besides his suggestion that damages involve quantifying the claimant’s reduction in welfare, Weinrib provides no explanation of what it means to compensate for factual losses. In particular, he does not explain whether compensation involves negating the factual loss, or providing a benefit to counterbalance the factual loss.

Robert Goodin comes closer to this point with his distinction between means-replacing and ends-displacing compensation.⁸ The former seeks to give the victim equivalent means for pursuing the same ends, while the latter seeks to help her pursue other ends in a way that leaves her equally well off. Giving an amputee an artificial limb is an example of means-replacing compensation because it provides her with equivalent means to pursue the

⁶ Weinrib, *The Idea of Private Law*, p. 40.

⁷ *Ibid.* p. 131.

⁸ Robert Goodin, “Theories of Compensation”, *The Oxford Journal of Legal Studies*, 9(1), (1989), 56 – 75 at p. 60.

kind of ends that require the use of legs. Giving the amputee a holiday to the Mediterranean, on the other hand, is an example of ends-displacing compensation because it allows her to pursue different ends, in an attempt to render her equally well off.

This distinction is similar to that between negating and counterbalancing, but there is an important difference. Giving a victim equivalent means will not always render her equally well off. An amputee who is given a prosthetic limb might be psychologically affected by the change in her appearance or self-esteem, even if she is able to pursue the same ends. Thus means-replacing compensation does not address all elements of the harm. Moreover, not all corrective measures can be understood as means-replacing or ends-displacing. Compensation for pain and suffering responds to a decrease in the victim's wellbeing, independently of how far her pain reduces her ability to pursue her ends, and this is better captured with the broader notions of negating and counterbalancing.⁹

2. Negating and Counterbalancing

A number of concepts are often used to describe the actions of corrective duties. Compensating, correcting, repairing, counterbalancing, removing and negating a state may be interpreted as equivalent descriptions of the same corrective action. Although 'correcting' and 'repairing' seem to refer to corrective actions in the most general sense, the terms 'negating' and 'counterbalancing' refer to two distinct and fundamental types of corrective

⁹ A similar criticism can be made of Arthur Ripstein's contention that "the law of torts protects each person's means against other persons". See Arthur Ripstein, "As If It Had Never Happened", *William and Mary Law Review*, 48, (2007), 1957 – 1999 at p. 1966. We can add that, whatever its merits as a legal insight, it is clearly unsatisfactory as a normative view. If D wrongfully injures V, causing her severe pain but without interfering with her ability or means to pursue her ends, D still owes V a corrective duty.

measure. The distinction can be illustrated with some simple examples. Consider the following cases, and let us assume that D is liable for V's harm in each of them.

Case 1: D injures V by breaking her leg, but because V is already under a general anaesthetic she suffers no pain or discomfort. An operation can fully repair V's leg before she regains consciousness.

Case 2: D injures V by breaking her leg, causing her pain and discomfort. An operation will fully repair V's leg.

Case 3: D injures V by breaking her leg, causing her pain and discomfort. The most effective operation on V's leg will render the limb usable but will not fully remove the pain and discomfort. There is also no painkilling medicine that will fully remove the pain and discomfort.

Case 4: D injures V by breaking her leg. No operation or treatment can restore the use of her leg or remove any of the pain and discomfort.

In case 1 the operation will negate the threat to V's wellbeing posed by the damage to her leg. If the leg is fully repaired before V wakes up, her wellbeing will not be affected by the injury. In this case a negatory duty is sufficient as there is no other detrimental state that must be counterbalanced. Given that V is under anaesthetic, the damage to her leg does not cause her reduced functionality or physical pain (assuming, of course, that the repair won't cause her additional pain when she wakes up). In case 2, the operation will only negate a *portion* of V's harm. It will fully restore the functionality to V's leg and prevent any further suffering. But it is impossible to negate the pain and discomfort V suffered before the operation, so instead D has a duty to provide some benefit to counterbalance the harm V has already suffered. Case 3 is similar to case 2 in the sense that D can only partially negate V's

harm. But in this case, D must provide a benefit to counterbalance the pain and discomfort V suffered before the operation *and* the future pain which medical treatment cannot prevent. Finally, in case 4, none of V's harm can be negated; D can only provide a benefit to counterbalance it.

It should be evident from these examples that many cases involve both negatory and counterbalancing duties. The content of the duty to counterbalance is largely determined by the possibility of negating the relevant states. Some detrimental states, most obviously harms such as physical disabilities and chronic pain, are ongoing. The longer the defendant fails to negate them the more harm they cause the victim, and the more harm must be counterbalanced. In cases such as 2 and 3, D is not able fully to negate the harm, and so she is unable to prevent V from suffering further. D instead has a duty to provide a benefit to counterbalance the expected future harm that she cannot now prevent.

I have introduced the concepts of negating and counterbalancing on an intuitive level, but we can now render them more precise. Negating a detrimental state, S, which an agent, V, is in, involves rendering V's future wellbeing identical to what it would be if not for S (strictly speaking, it is the properties that determine V's wellbeing that are equal or identical, rather than V's wellbeing itself, but for convenience I will continue to use the shorter formulation).¹⁰ This does not just mean ensuring that V has the same *amount* of

¹⁰ To simplify, I assume throughout this paper that the victim's wellbeing would remain constant if not for the detrimental state. This will not be true in some cases, such as those involving preemptive harms, where the harm inflicted on a victim by one agent prevents an equivalent or greater harm being inflicted on the same victim by another. In these cases, V's wellbeing would not have remained constant had the first agent not inflicted the harm. It is important that we have the theoretical resources to deal with these cases, as they pose difficulties for the measurement of setbacks to wellbeing. However, since my main interest is in the distinction between negating and counterbalancing, which can be investigated independently of these difficult cases, I leave them to one side.

future wellbeing that she would have if not for S. Rather, it means ensuring that V's future wellbeing is of the same *shape* and *quality* that it would be if not for S. Conversely, counterbalancing a state, S, involves rendering V's future wellbeing equal to what it would be if not for S, even if it differs in shape and quality.¹¹

Let us begin with the distinction between the shape and amount of V's wellbeing. Consider the following simplified example, represented in figure 1 below. Suppose V suffers a broken leg at T2, and as a result her wellbeing decreases. V receives a benefit which temporarily increases her wellbeing, after which it reverts to the level it was prior to the injury. We can further assume, as depicted in figure 1, that the *increase* in wellbeing V enjoys as a result of the benefit relative to her wellbeing before the injury is equal to the *decrease* in wellbeing V suffers as a result of the injury relative to her previous level of wellbeing. We also make the simplifying assumption that V's wellbeing would have remained constant if not for the injury. In this case, the benefit counterbalances the decrease in wellbeing because V ends up with the same *amount* of wellbeing that she would have had if not for the injury at T2. But although V's average wellbeing is the same, the *shape* of her wellbeing through time is not. It has been decreased and then increased rather than remaining constant. Conversely, negating a state makes the future shape of the agent's wellbeing identical. In case 1, for example, represented in figure 2 below, where D breaks

¹¹ I assume that there are cases in which a person's wellbeing can be rendered roughly equal overall by providing them with some benefit. Wellbeing is, of course, a complex and contentious concept. The assumption that comparisons of wellbeing can be made might also be controversial, but the possibility of rough equality is sufficient to motivate arguments about negating and counterbalancing (for more on this possibility, see Ruth Chang, "The Possibility of Parity", *Ethics*, 112(4), (2002), 659-688 and Derek Parfit, *Reasons and Persons*, p. 481). Moreover, when the currency of a corrective duty can be quantified precisely, as in cases involving economic losses, rendering a person identically well off and equally well off with respect to those resources can be clearly distinguished.

V's leg at T2 and repairs it before V regains consciousness, the shape of V's future wellbeing is identical to what it would be if not for the injury, because it simply remains constant.

Figure 1

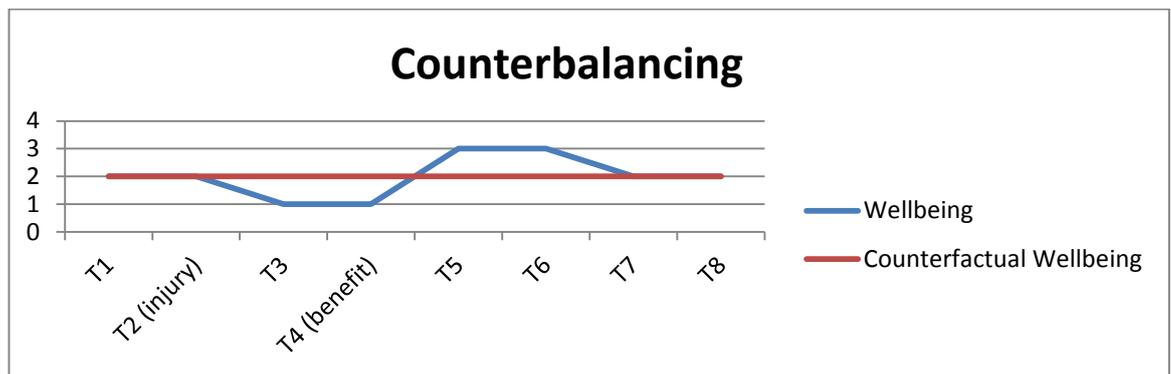
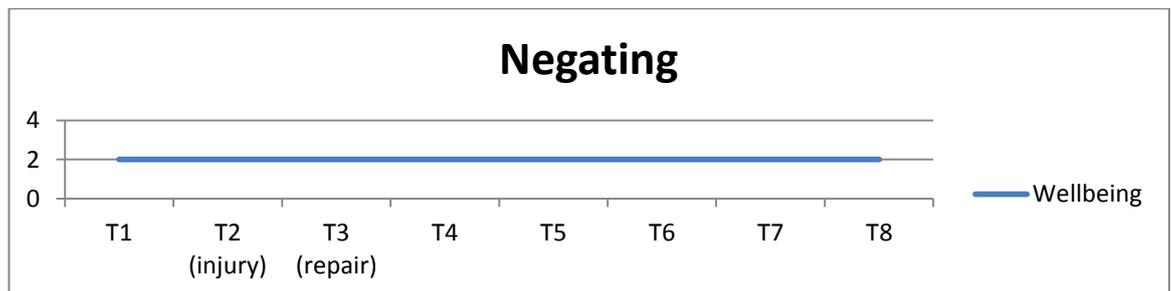


Figure 2



Negating also differs from counterbalancing because it renders V's future wellbeing identical in terms of *quality*. The above figures are simplified because they do not distinguish between separate components of wellbeing. When V suffers a broken leg, there are a number of ways that her wellbeing decreases, including her experience of physical pain, her reduced ability to pursue her projects, the effect of the injury on her self-esteem, and so on. In the counterbalancing case, when V is given a benefit at T4, even if we suppose her overall wellbeing is equalised, it may differ in quality. For example, if the benefit is a sum of money with which V builds a library in her house, the components of her wellbeing will be different compared to those she would have enjoyed if the injury had not occurred.

Had the injury not occurred, she might have continued gaining satisfaction from her work. Although her injury excludes *this* component of wellbeing, she is equally well off because she is able to spend time in her library, which represents a different component. Conversely, in the negation case, when D negates the damage to V's leg while she is still unconscious, V's future wellbeing is not only equal in amount, but also identical in quality, with what it would have been had the injury not occurred.

Negating, then, renders the victim's future wellbeing identical to what it would be if not for the harm, in terms of amount, shape and quality. Counterbalancing renders her future wellbeing equal to what it would be if not for the harm only in terms of amount. With the distinction clarified, we can now address the reasons that should motivate us to accept its significance.

3. Primary and Secondary Duties

A. Preventive and Non-Preventive, Forward and Backward Looking.

The first reason why we ought to accept the distinction is that it allows us to clarify the debate about whether it is justifiable to impose secondary duties on those who have breached their primary duty to avoid committing a wrong. In order to explain this, I must introduce some further terminology. Duties whose function is, at least in part, to *avoid* some state are *preventive duties*. Preventive duties can be divided into two groups: *first-person preventive* and *third-person preventive*. First-person preventive duties require an agent to avoid *causing* some state. Third-person preventive duties require an agent to act so that some state is avoided, even if that agent would not have caused it. Duties not to harm or commit wrongs are first-person preventive as they require an agent to avoid causing one of these detrimental states. A duty to save a drowning child (which one is not causally responsible for

endangering) is third-person preventive as it is designed to avoid an outcome, but one that the duty-bound agent would not have caused.

Duties are *forward-looking* if they are directed at some state that will occur in the future or has already begun to occur but will continue into the future. So if D punches V in the face, causing her some transient pain but not placing her in an ongoing harmed state, any corrective duty D owes will be backward-looking. If the punch breaks V's nose, placing her in an ongoing harmed state that requires some intervention, a duty to repair V's nose will be forward-looking as it responds to a state that continues into the future. Alternatively, if D exposes V to noxious dust that will cause her to develop cancer in five years, a duty in response to this will be forward-looking as development of the cancer will be a new state beginning at a future date. As before, we can also distinguish between forward-looking duties that are directed at states for which the duty-bound agent is *causally* responsible, and those that are directed at states for which she is not.

All preventive duties are forward-looking, since it is impossible to avoid a past state. However, not all forward-looking duties are preventive. If a future state cannot be avoided there can be no preventive duty requiring some agent to avoid it, but there may still be a forward-looking duty to do something else in response to the unavoidable future state, as there is in the case where V's exposure to noxious dust will cause her to develop cancer. In this case D may have a duty to give V a different good, such as a sum of money, in lieu of averting the threat. Primary duties are both preventive and forward-looking. They are preventive as they require us to *avoid* committing wrongs, and since they are preventive they are also forward-looking. Negatory duties are *also* preventive since they require us to avoid the continuation of some detrimental state by negating it, and since they are preventive they are also forward-looking. Counterbalancing duties are *never* preventive as they do not seek avoid harm. But they can be forward-looking since they may seek to counterbalance a future state that cannot be avoided. They may also be backward-looking if they seek to counterbalance a past state.

The close connection between primary and negatory duties is notable. Suppose D wrongfully causes V to be in a specific harmed state, H, at T1. She cannot now fulfil her primary duty not to put V in H, but she can fulfil a very similar one. She can fulfil a duty not to cause V to be in H at T2, where T2 is the earliest point at which she is able to negate the effects of her wrong. If she fails to do this, she will have caused V to be in H at T2. The only difference between the descriptions of these duties is temporal. Both duties are first-person preventive as they require D to avoid causing V to be in H, albeit at different times. We might therefore think of negatory duties as instances of general duties to avoid harming others. Similarly, we might think of them as instances of general duties not to wrong others, since once a person suffers a wrongful harm the wrongdoer commits a further wrong if she fails to fulfil her corrective duties. Since, unlike primary and negatory duties, counterbalancing duties do not prevent detrimental states, the latter cannot be seen as instances of duties to avoid harms or wrongs. It is only when a harm is irrevocable or unavoidable that a benefit must be provided to counterbalance it. But as noted above, although counterbalancing duties are not preventive, they may, and often do, respond to *future* harm and so are forward-looking.

B. The Justifiability of Secondary Duties

With these details in place, we can identify a few confusions inherent in the debate about whether imposing secondary duties is justifiable. The breach of a primary duty is often held to *ground* a secondary duty. According to these views, imposing such a duty is the next best way to uphold the right that was violated, so the justification is ‘carried over’ from the primary duty. For example, Stevens argues that “the secondary obligation to pay money imposed upon the wrongdoer can be seen as the law's attempt to reach the ‘next best’

position to the wrong not having been committed by him in the first place”.¹² Similarly, Weinrib claims that “with the materialization of wrongful injury, the only way the defendant can discharge his or her obligation respecting the plaintiff’s right is to undo the effects of the breach of duty”.¹³ As John Gardner has suggested, these views should not describe the initial breach as a mere *condition* of a secondary duty, as this fails to capture the justificatory relationship between them. A secondary duty is imposed if *and because* a primary duty has been breached – the breach of a primary duty is often necessary and sufficient to justify a secondary duty.¹⁴

The assumption that a breach is *necessary* appears to follow from the fact that primary and secondary duties are fundamentally distinct, albeit closely related. Otherwise, secondary duties might need no greater justification than primary duties. But as indicated above, one category of secondary duties – negatory duties – shares some properties with primary duties. More specifically, many primary and negatory duties are first-personal and preventive. They are designed to prevent either the initial occurrence or the continuation of a detrimental state caused by an agent. Perhaps, therefore, it is sufficient to ground a negatory duty that, having caused V to be in H at T1, D will cause V to be in H at T2 unless she removes H. Since V’s harmed state is ongoing, only some of the overall costs have been incurred, and the potential costs may ground a negatory duty, the fulfilment of which will avert those costs, in the same way they ground a primary duty to avoid causing V to be in H in the first place. This depends on what role the preventiveness of primary and negatory duties plays in their justification.

¹² Robert Stevens, *Torts and Rights*, p. 59. See also Arthur Ripstein, “As If It Had Never Happened”, p. 1968

¹³ Ernest Weinrib, *The Idea of Private Law*, p. 135.

¹⁴ See John Gardner, “Backwards and Forwards with Tort Law” in *Law and Social Justice*, Eds., J. Keim-Campbell, M. O'Rourke & D. Shier, MIT Press, (2005), pp. 24 – 31.

I defined a secondary duty as one that prescribes a corrective measure in response to a detrimental state caused by the breach of a primary duty. Negatory duties sometimes fit this description, so on this model they are classed as secondary. The truth, however, is that they do not fit simply into this binary scheme. The properties of these duties with potential moral relevance are not uniformly distributed across all those labelled 'secondary'. This is a problem, since the debate about their justification is framed in terms of this scheme. The key question is really this: precisely which features of secondary duties are thought to pose a special justificatory problem? The fact that primary duties are preventive cannot explain why negatory duties require special justification, since these are also preventive. Another thought is that whilst all primary duties are forward-looking, secondary duties are backward-looking. But again, this is a misunderstanding. Negatory duties, like primary duties, are preventive and therefore forward-looking. Counterbalancing duties, though not preventive, *can* be forward looking if they are designed to counterbalance an event that will unavoidably occur in the future, as in the case where a victim develops latent cancer.

To put the point another way, if imposing preventive duties is more justifiable, *ceteris paribus*, than non-preventive duties, then imposing both primary and negatory duties is more justifiable than counterbalancing duties. If imposing forward-looking duties is more justifiable, *ceteris paribus*, than backward-looking duties, then imposing primary, negatory and forward-looking counterbalancing duties is more justifiable than backward-looking counterbalancing duties. Neither of these distinctions – between preventive and non-preventive, and forward-looking and backward-looking duties – explains why, on the traditional understanding, secondary duties require special justification.

There *are*, however, two important differences between primary duties on the one hand and negatory and counterbalancing duties on the other. First, it is often more *burdensome* to fulfil the latter than the former. Most wrongs are avoidable without a significant degree of effort. One can avoid breaching a duty of care in negligence by using a reasonable degree of care and skill. If this duty is breached, the corrective obligations that follow may be much more costly to fulfil. If the victim suffers a broken arm and chronic

discomfort, repairing the break and counterbalancing the discomfort may cost the wrongdoer substantially more than it would have cost to meet her standard of care. It must be remembered, however, that this is *typically* rather than necessarily true: we can imagine a careless billionaire for whom counterbalancing harm *ex post* is less burdensome than meeting a reasonable standard of conduct *ex ante*.

Second, primary duties are generally negative whereas negatory and counterbalancing duties are positive. In other words, primary obligations do not usually impose ends on the duty-bearer – they merely require her to refrain from performing some act, X, leaving her free to do anything but X. Corrective duties require the duty-bearer to set as her end the negation or counterbalancing of some detrimental state. Perhaps these two features explain the way the justificatory issue has traditionally been framed.

I have outlined four separate features of corrective duties: preventiveness, temporal orientation, costliness and the distinction between negative and positive duties. All of these features may have some moral significance, although assessing this is a difficult normative inquiry that I have not attempted to undertake. The point I want to reiterate is that these potentially significant properties are not uniformly distributed between negatory and counterbalancing duties. Referring generally to corrective duties will obscure this important fact, so recognising the distinction is a prerequisite to assessing the normative significance of these duties.

4. What Corrective Duties May Exist?

The second reason we ought to accept the distinction is that it affords a more nuanced set of possibilities about which corrective duties *may* exist. So far I have focussed on those that follow from wrongful positive acts, but I noted that the analysis of such duties is not conceptually bound to these. They may arise by virtue of wrongful omissions, in addition to

other events besides wrongs, such as causing harm without fault. Some may even arise without wrongdoing *or* causation.

Consider the following cases.

Case 1: D knocks over V's fence. If it remains this way V will suffer a loss. D can put up the fence with minimal cost.

Case 2: D sees that V's fence has fallen over. If it remains this way V will suffer a loss. D can put up the fence with minimal cost.

Case 3: D knocks over V's fence and V's chickens escape, causing her a loss. D can give V a benefit that equalises her wellbeing which will involve a minimal cost for D.

Case 4: D sees that V's fence has fallen over and V's chickens have escaped, causing her a loss. D can give V a benefit that equalises her wellbeing which will involve a minimal cost for D.

The corrective measures specified in cases 1 and 3 are negatory as they involve removing a threat of future loss to V. Putting up the fence will render V's future wellbeing identical both in terms of its shape and quality. The corrective measures specified cases 2 and 4 involve counterbalancing a loss that has already occurred. Given that V's wellbeing has already been reduced, it cannot be rendered identical, but by providing a benefit, D can render her equally well off. Moreover, in cases 1 and 3, D *causes* the detrimental state, whereas in cases 2 and 4 there is no causal link between D's conduct and the detrimental state. Finally, let us suppose that D is not at fault for knocking over V's fence in cases 1 and 3.

Many people would accept that D has a corrective duty in case 1 and not in case 4. If D knocks over V's fence through her voluntary action, even if she is not at fault, she ought to put it up if doing so involves minimal cost. Conversely, in case 4 it is implausible to

require D, a passer-by, to pay for a loss that she did not cause. The loss has already been incurred, and (let us imagine) D is no better situated than anyone else to counterbalance it, so there seems to be no reason why the burden should fall on D. Cases 2 and 3 may be more controversial. In case 3 D *causes* a loss through her voluntary action, but she was not at fault for doing so and is not culpable. The plausibility of imposing liability on her may turn on the opportunity she had to avoid causing the loss. In case 2 D does not cause any loss, but the forthcoming loss is preventable and she is the only person who has the opportunity to prevent it. Views about whether corrective duties exist in these cases will turn on the significance of preventability and causing loss through voluntary action.

It is not my present purpose to explore each of these questions. The point I want to make is this: it is more plausible that there is a corrective duty in case 2 than case 4, even though D plays no role in knocking over V's fence in either case. A number of features of case 2 count in favour of D owing a corrective duty: (1) the loss is preventable (2) D is the only person who can negate the threat (3) D is able to do so with minimal cost. Only (3) is also true of case 4, so it is less plausible that a corrective duty exists there. At the very least, in case 4 there seems to be no reason why V should be the *only* person who is subject to a corrective duty, given that she is no better situated than anyone else to counterbalance the loss. If loss spreading is not possible, it is implausible that D has any duty in case 4 at all. In case 2, where loss spreading is also impossible, there is still a reason to impose a duty on D.

This shows that negatory duties may arise in the absence of either wrongdoing or causing loss without fault in cases where the same is not true of counterbalancing duties. Of course, it may still be true that in *other* cases, counterbalancing duties can arise without wrongdoing or causation, such as political duties to contribute to a tax scheme designed to counterbalance the disadvantages of those with congenital disabilities. The fact that in case 2 D is uniquely well situated to avoid an interference in V's wellbeing seems to explain the intuitive difference between the cases. This suggests that the possibility of preventing some loss goes some way towards justifying a corrective duty when other key features, such as wrongdoing and causal responsibility, are absent.

I will not pursue these possibilities any further. My point is that the existence of a corrective duty in case 2 is an open question which is obscured unless we properly distinguish between negatory and counterbalancing duties.

5. Identically Off and as Well Off

The previous two arguments aimed to show that the distinction helps to properly frame normative questions by drawing attention to features of negating and counterbalancing that have potential moral relevance. The final argument differs from these as it demonstrates the actual significance of the distinction. We have moral reasons to care about the difference between being identically off and as well off, and therefore we have reasons to care about the distinction between negating and counterbalancing as corrective actions. The implications of this distinction for permissible action have been anticipated by other writers. Arthur Ripstein notes that “if money could make an injured person whole, it seems that injuring someone and then paying them is just as good as not injuring them at all”.¹⁵ That compensation is inadequate in this respect is a highly intuitive idea, but it has not been investigated in much detail. Figurative phrases such as ‘making whole’ require further analysis as they gloss over the distinction and fail to draw out its implications.

The argument will proceed as follows. I begin by analysing pairs of cases that reveal the moral difference between being identically off and as well off. These pairs must be designed so that this is the only operative difference between the two cases. For example, none of the cases should involve one person using another person or their property as a means, as this may have independent moral relevance which could obscure the effect of the distinction under investigation. Once the case analysis has established the relevance of the distinction with respect to permissible action, I provide an explanation for these cases based

¹⁵ Arthur Ripstein, “As If it Had Never Happened”, p. 1957.

on the idea that interfering with another person's wellbeing (as opposed to reducing it) is distinctively wrong.

The argument is motivated by a comparison between two types of case: those in which an agent renders someone identically off and acts permissibly, and those in which an agent renders someone as well off and act impermissibly. Consider the following two cases:

Case 1: V's fence is in D's way. D knows that if she knocks it down she will be able to put it back up before V's wellbeing is affected.

Case 2: V's fence is in D's way. D knows that if she knocks it down she will not be able to put it back up before V's chickens escape. But D can give V a benefit that will counterbalance the decrease in V's wellbeing caused by the escape of the chickens.

Let us further assume that in both cases D's reason to knock down V's fence is not important enough to allow D a defence of necessity. It is not required to save life or limb or avoid any other significant loss.¹⁶ Nevertheless, it is plausible that in case 1, D does not act wrongly if she knocks down V's fence. We can imagine that D has *some* reason to do it: perhaps she has dropped one of her possessions behind it and the only way to retrieve it is to pull the fence down. It seems that if V will be rendered identically off, it is permissible for D to do this. We need not, in fact, rely on as strong a claim as this. It is enough to say that even if D acts wrongly, the wrongness of this act has nothing to do with its effect on the shape of V's wellbeing. It might be thought that V is wronged simply because her property has been used

¹⁶ For a discussion of liability to compensate in cases where an infringement of a right is necessary to avert a significant loss, see Joel Feinberg, *Rights, Justice, and the Bounds of Liberty*, Princeton: Princeton University Press, (1980), p. 230 and Judith Jarvis Thomson, *Rights, Restitution, and Risk*, Cambridge: Harvard University Press, (1986), p. 71.

in a way to which she does not consent that is not justified on grounds of necessity. But on this view, even if the act is wrong, it is not wrong because of its effect on V's wellbeing.

In case 2, it is plausible that D acts wrongly if she knocks down the fence. She does not have a very good reason to knock it down, and doing so will affect V's wellbeing in a morally significant way, even though her average wellbeing will not be reduced. As before, we can rely on an alternative weaker claim. Even if D's act is permissible all things considered, it is sufficient that the way D affects the shape of V's wellbeing is a moral consideration that counts against D's act. This is enough to show that there is a moral difference between being identically off and as well off.

We assumed that D's reason to knock down V's fence in these cases was not very strong. This was because if D had a very good reason (to save a person's life, for instance) it would be permissible to do so in both cases. But the distinction is still operative in some cases that involve the permissibility of acts that save others from harm. Consider the following two cases:

Case 1: D wants to drag V1 from some wreckage. If D drags V1 from the wreckage, V1 will emerge unscathed, but if D does nothing then V1 will lose a finger. The only way D can drag V1 from the wreckage is by moving the unconscious body of V2, who is lying in the way. If D does this, V2 will suffer a broken leg, which will cause permanent damage unless treated quickly. Luckily, D has brought his home surgery kit and will be able to repair V2's leg before she wakes up, thus averting the threat to her wellbeing.

Case 2: Identical to case 1, except that D has forgotten his home surgery kit. If D moves V2's body to save V1's finger, V2 will live the rest of her life with a damaged leg, and will be unable to pursue many of the projects she currently values. V2 also has a genetic mutation that will cause her to develop a condition in the future involving chronic pain. To counterbalance the damage to her leg, D can give V2 a new medical treatment, too rare to be

otherwise available to her, which will dramatically alleviate the symptoms of this condition when they appear.

In case 1, it is plausible that D is permitted to move V2. It is true that D must damage V2's body to do this, but D can render her identically well off, and the need to save V1's finger seems to justify this. In case 2, however, it seems implausible to think that D is permitted to move V2's body, even though she will ultimately be equally well off. Of course, we must assume here that V2 *will* be as well off. D inflicts one type of harm (breaking her leg) and then prevents another type of harm (chronic pain), and *ex hypothesi* these harms constitute a roughly equal reduction in V2's wellbeing. The harms involved can be adjusted to suit contrasting views about the extent to which different harms setback a person's wellbeing. My claim is that the intuitive difference between these two cases will survive these variations.

What explains these results? We can explain them in terms of the *interference* with the victim's life when she is rendered equally well off, which is not present when she is rendered identically off. We can pinpoint three distinct features of interference. One feature is that making a victim as well off, unlike making her identically off, sometimes reduces her ability to set her ends, and this is morally significant even if the benefit equalises her wellbeing. In case 2, although the chronic pain reduces V2's wellbeing to the same extent as the broken leg, it may not equally reduce her ability to set her ends. A permanently damaged leg rules out a broad range of ends that involve the use of one's leg's in a way that chronic pain does not, and may therefore restrict her end-setting capacity to a greater degree.

A second and closely related reason is that it is also important to pursue the specific ends that one has set. That this is distinct from one's *capacity* to set one's ends is clarified by the following example. Suppose D injures V, reducing her ability to set her ends. To counterbalance this, D gives her a benefit that increases her ability to set her ends, such as a large sum of money. In this case, it is not true that D has reduced V's average ability to pursue her ends. Nevertheless, D may have affected the specific ends V is able to pursue.

We can imagine that as a result of her injury, V is no longer able to play her favourite sports. This is an interference with V's life even though she now has enough money to pursue a variety of other ends, and thus her end-setting capacity has not been reduced.

Finally, even if D gives V a benefit that equalises her capacity to set her ends *and* allows her to pursue the ends she had previously set, D still interferes with V's life by affecting the *time* at which she can pursue them. Suppose that in case 2, V2's injuries will prevent her from pursuing a set of ends, S, between T1 and T2, a period of five years. After this period, the damage to her leg will be repaired. However, at T2, V2's ability to pursue S will be removed anyway by her genetic medical condition until T3, also a period of five years. As before, D can give V2 a benefit that will treat her genetic medical condition, enabling her to pursue S between T2 and T3. In this case, D has not reduced her overall capacity to set ends *or* her ability to pursue the ends she has set, but D has still interfered with the time at which she can pursue her ends. This interference also plausibly counts against D's actions in case 2.

Armed with the notion of interference, we can show how the distinction between negating and counterbalancing has a number of implications for contemporary tort theory. First, it supports and refines John Gardner's defence of the moral norm of corrective justice. Gardner claims that "once the time for performance of a primary obligation is past, so that it can no longer be performed, one can often nevertheless still contribute to satisfaction of some or all of the reasons that added up to make the action obligatory. Those reasons... call for next-best satisfaction, the closest to full satisfaction that is still available".¹⁷ He refers to this as the "continuity thesis".

We can now see that there are at least two sets of reasons to avoid harming others. One set is harm-based (grounded in the need to avoid reducing the wellbeing of others) and the other is interference-based. When a primary duty is breached, the set of reasons that subsist point directly to a negatory duty. This is because, assuming that at least some

¹⁷ John Gardner, "What is Tort Law For?", *Law and Philosophy*, 30, (2011), 1 – 50, at p. 33.

elements of the harm can be negated, the interference-based reasons will still be operative, and the only way to satisfy them is to negate as much of the harm as possible. Only when this possibility has been exhausted will the interference-based reasons disappear (because ought implies can) leaving only the harm-based reasons. These reasons will then support a purely counterbalancing duty, designed to render the victim equally well off. This sketch shows how the continuity thesis and the present distinction are mutually supportive.

Second, recall Gardner's comments that it is surely legitimate for tort law to aim at deterring people from committing torts by doing corrective justice. "Wouldn't it be better," he asks, "even from the perspective of tort law itself, if there were less correcting to do thanks to the fact that fewer legally recognized wrongs, fewer torts, had been committed?"¹⁸ We noted in the introduction that the legitimacy of this instrumentalist goal may have further implications. It may be used as a decision mechanism when resources are limited. If there are only resources to enforce one of two corrective duties, it is permissible to enforce the one that will have the greatest deterrent effect. Such a theory would allow instrumentalist goals to be pursued subject to certain conditions. The need to minimise or eliminate interference shows why avoiding harm is preferable to fully counterbalancing it, and hence why deterrence is a legitimate aim of tort law. It is this analysis that explains Gardner's intuitive comment that deterring torts is surely better from the perspective of tort law itself.

Third, the concept of interference, as it is defined here, gives normative strength and clarity to the rejection of the economic view of tort liability as a practice of pricing harmful behaviour. To adopt the terminology of economic theorists, 'liability rules' (rules that inform people 'if you X, you will pay Y') differ from 'property rules' (rules that tell people 'don't/do X') because they ignore the importance of interference. The cases considered in this section suggest that certain actions are impermissible despite the availability of full

¹⁸ John Gardner, "What is Tort Law For?", p. 23.

compensation, because they interfere with the lives of victims. This is why liability, in these and many similar cases, ought not to take the form of a ‘liability rule’.¹⁹

It is worth noting here that similar arguments can be made about interference if we adopt other accounts of the currency of corrective duties. I mentioned at the outset that although I discuss cases of setback or threat to wellbeing, other currencies could be adopted. For example, we could adopt the hypothetical preference account, which initially seems to be more autonomy-preserving than the wellbeing account. Suppose that a victim would be indifferent between (1) suffering internal injuries and going on a cruise and (2) not suffering the injuries. If preferences are the correct currency of counterbalancing, the cruise will adequately counterbalance V’s injuries in this case. But this does not eliminate the interference in V’s life and does not make it permissible for D to interfere with V in similar cases to those discussed above. D is still affecting her capacity to set ends, the specific ends she can pursue and the time at which she can pursue them. It must be remembered, however, that I am referring to V’s hypothetical preferences. It remains true that D could permissibly interfere with V if she consents to this interference.

6. Negating, Counterbalancing and Remedies in Private Law

Having established the principled difference between negating and counterbalancing, I will now outline some ways in which the distinction can help us explain and justify specific legal remedies and practices. First, it helps to explain why it is typically not acceptable, both in principle and in practice, for a wrongdoer to delay fulfilling her corrective duties. This is not

¹⁹ On the distinction between ‘liability rules’ and ‘property rules’ see Guido Calabresi, *The Costs of Accidents*, Yale University Press: New Haven and London, (1970) at pp. 136 – 7 and Guido Calabresi and Douglas Melamed, “Property Rules, Liability Rules and Inalienability: One View of the Cathedral”, *Harvard Law Review*, 85, (1972), 1089 – 1129.

impermissible merely because the victim will suffer more harm in the meantime, since we can imagine a wrongdoer who is prepared to fully counterbalance the harm when she is good and ready. If the wrongdoer's corrective obligations consist solely in the duty to counterbalance, there would be nothing wrong with this behaviour. In either case the victim will end up as well off as she would be if not for the harm. We could simply build in to the duty to counterbalance the additional requirement that the benefit must be conferred as soon as possible, but this move is purely arbitrary. This requirement is, however, a natural feature of the duty to negate. This is because the duty is preventive. Its purpose is to stop the victim suffering further harm, so it is important that it be fulfilled as soon as possible, as the sooner it is fulfilled the more harm is likely to be prevented. In contrast, the criteria for counterbalancing can often be fulfilled regardless of the magnitude of the harm. The normative priority of negating over counterbalancing helps to explain why a wrongdoer cannot delay fulfilling her corrective duties once liability has been established.²⁰

Our discussion also helps to explain two common remedies in private law. First, injunctions have a clear negatory purpose. If a defendant commits an ongoing wrong, a successful injunction will negate its continuation. Injunctions, like negatory duties, are therefore preventive. An injunction may also be granted where the defendant intends to commit a wrong in the future. The injunction can be interpreted as imposing a negatory duty on the defendant to avert the threat of committing the wrong posed by their intention to do so. In this case, as I mentioned in the discussion of primary duties, the negatory duty will be effectively equivalent to a primary duty to avoid committing the wrong. That is, they are both first-person preventive duties not to commit a particular wrong.

Similarly, a mandatory injunction or an order of specific performance for breach of contract is sometimes preferable to monetary damages because it allows the victim to pursue

²⁰ Admittedly, this is not the whole story. If negating is impossible, there might still be some reason to counterbalance earlier rather than later. Above I have identified one reason why corrective duties have implicit time requirements, but I do not preclude the possibility that there are others.

the same end or ends she would have pursued if not for the wrong. It therefore reduces or removes one type of interference that makes counterbalancing inferior to negating. It improves the victim's ability to pursue the specific ends that she has set. If the same result cannot be achieved through damages, injunctions may be preferable as they are closer to pure negation.

Moreover, counterbalancing in a way that does not reduce the victim's capacity to set her ends is preferable to counterbalancing in a way that does. This is one reason why instrumental goods such as monetary awards are often the best way of counterbalancing setbacks to wellbeing. This yields a principled reason why money is an appropriate form of compensation, beyond the practical advantage that it is the easiest to transfer. Hence, if a defendant can do one of two things: (1) give V a sum of money that enhances her capacity to set her ends and gives her the opportunity to render herself equally well off or (2) send V on a cruise that will render her equally well off, the need to maintain V's capacity to set her own ends explains why (1) is a preferable remedy.

As I have mentioned, throughout this paper I have only discussed negating and counterbalancing setbacks or threats to *wellbeing*. But in cases where the victim's past or future losses are economic, resources rather than wellbeing are the appropriate currency, and some cases may involve a plurality of currencies. For example, if D wrongfully causes V a physical injury, as a result of which she is unable to work for five years, V suffers both a setback to wellbeing and resources. Her loss of resources does not necessarily setback her wellbeing, since she might have spent the money on a gambling addiction that would make her worse off overall. The negatory and counterbalancing duties owed by D may be directed at different currencies, even though they are both fulfilled by transferring damages. The setback to V's wellbeing is counterbalanced by a sum of money that gives her the opportunity to equalise her wellbeing, and the threat to her resources is negated by the receipt of identical resources. We can understand cases in which damages compensate for loss of earnings as implementing negatory duties, applied to resources rather than directly to wellbeing.

Another implication is that a sum of damages can fulfil both negatory and counterbalancing duties. This is obscured by the fact that an award of compensation is usually interpreted as the fulfilment of a single corrective duty. It would also be wrong to think of damages as necessarily a backwards-looking remedy. Return to the case where D exposes V to noxious dust, causing her to develop cancer five years after the exposure. Assume also that medical treatment cannot prevent the onset of the cancer. If this can be proven, V may seek damages before the injury materialises, in which cases the compensatory award will be non-preventive but forward looking.

7. Choice and the Normative Priority of Negating Over Counterbalancing

I have provided arguments to show that, *ceterus paribus*, negating has normative priority over counterbalancing. I will conclude by briefly outlining cases where this relationship does not hold. These exceptions to the priority rule do not give rise to any inconsistency since they stem from the same principle that establishes the priority rule itself; namely, the distinctive wrongness of interference.

To recap, here is an example of the normative priority of negating over counterbalancing. Suppose that D negligently injures V, and to fulfil her corrective duties, D can either negate or counterbalance the harm. Suppose further that negating the harm is more burdensome for D than counterbalancing it. Perhaps negating requires an operation, costing £5000, while a Mediterranean cruise costing £4000 would be enough to render V equally well off. The previous arguments suggest that D has a duty to pay for the operation and cannot complain that V could be made equally well off at a cheaper cost to D.

One problem with the priority rule is not based on concern for interfering with the *victim*, but rather a reluctance to interfere with the defendant. One issue that ought to be mentioned, though there is no space to discuss it in detail, is that in the common law of contracts an award of damages will often be expressly favoured by the courts over an order

of specific performance. This looks like a preference for counterbalancing over negatory duties because an award of damages interferes with a defendant's ability to set her ends less than compelling performance of her contractual obligations. This may also help to explain why prohibitory injunctions are granted more readily than remedies of specific performance,²¹ since specific performance forces a defendant to adopt a particular end while a prohibitory injunction commands her not to interfere with the ends of others.

Moreover, the priority rule can be reversed by the *actual* choices or preferences of the victim. In the above case, suppose that (1) negating and counterbalancing are equally burdensome for D and (2) V *prefers* counterbalancing over negating. In other words, V would rather have the cruise than the operation. Here, counterbalancing is Pareto optimal and probably permissible. It would be odd to insist that V's harm should be negated when she in fact prefers an alternative benefit which is not more burdensome for the defendant. Moreover, if D gives V a sum of money, she may be able to obtain a good that she prefers to the operation *and* the cruise. This good need not directly contribute to V's wellbeing. V may instead spend her damages on improving her daughter's education. If so, her harm will not be counterbalanced (she is not rendered equally well off) because she forgoes this possibility to further the interests of her daughter. Another conclusion we can draw is that it is unnecessary for D to *ensure* that V's harm has been negated or counterbalanced. It is sufficient if D provides V with the opportunity to fulfil the appropriate corrective action, leaving the decision about which particular good to pursue to V. This enhances the autonomy of the victim by giving her a broader range of choices about which ends to pursue.

8. Conclusion

²¹ See the judgement of Lord Cairns, L. C. in *Doherty v Allman* (1878) 3 App. Cas. 709, 719 – 20.

Let me summarise the previous discussion. I began by noting that although some scholars have distinguished between different types of corrective measures, there has been no acknowledgment of the distinction between what I call negatory and counterbalancing actions. These actions may apply to different currencies, including resources, wellbeing and so on. Using wellbeing as an illustration, I defined negation as rendering the victim's future wellbeing identical to what it would otherwise be, and counterbalancing as rendering the victim's future wellbeing equal to what it would otherwise be. Equalising the victim's wellbeing does not imply that it has the same distribution through time that it would otherwise have, or that it consists of the same components.

I then gave three reasons why we ought to accept the significance of the distinction. The first is that it helps to clarify debates about the justifiability of imposing secondary duties. The idea that secondary duties require special justification suggests that they are distinct from primary duties in some fundamental way. But labelling corrective duties 'secondary' conceals the fact that negatory duties share a basic property with primary duties: they are both preventive. While primary duties usually aim to prevent the occurrence of some detrimental state, negation aims to prevent its continuation. Preventiveness is one property with potential moral relevance that is not uniformly distributed across all secondary duties. The second and closely related reason is that recognising the distinction opens up important normative questions. For example, if preventiveness is indeed a significant property, negatory duties may exist, *ceteris paribus*, where counterbalancing duties do not. Ultimately, we must distinguish more accurately between different properties of corrective duties and investigate their moral significance, to determine whether they are justified and if so which we should enforce.

The third argument is perhaps the most important. Whereas the first two emphasise the value of the distinction for clearly framing normative questions, the third seeks to establish its actual moral significance. The argument is essentially that we have reasons to care about the difference between being identically off and as well off, which are explained by the notion of interference. Negating a detrimental state eliminates the future interference

of that state with the victim's life, whereas this is not true of counterbalancing. This moral difference has implications for permissible action. We looked at cases in which the prospect of interfering with the victim renders certain acts impermissible even though she is rendered equally well off. Conversely, these acts would be permissible if the prospect of interference is removed.

I then argued that the distinction can help explain and justify certain practices and remedies in private law. In particular, I argued that one feature of interference – its impact on the victim's ability to set her own ends – gives principled reasons in favour of a number of remedies. First, injunctions, which serve a clear negatory purpose, prevent others from interfering with the victim's ability to pursue an end she has set. Second, monetary compensation minimises interference because it has instrumental value, giving the victim a greater choice about which ends to pursue.

Finally, despite the fact that there are moral reasons to prefer negating over counterbalancing, the preferences of the victim can modify this priority rule. A victim may choose to be made equally well off even if it is possible to make her identically off; she may even choose to benefit another person rather than be made equally well off. This is simply another way to safeguard the victim's ability to set her own ends.

Actionability

Actionability is concerned with whether the consequences of wrongdoing, or non-wrongful conduct that incurs liability, are compensable or subject to other corrective actions. Even if it is established that A wrongs B, it is a further question whether the particular consequences of the wrong merit compensation. There are a cluster of principles governing legal actionability. In negligence, perhaps the most famous is the ‘remoteness of damage’ rule, which limits actionable losses to those that are reasonably foreseeable. Another is the ‘thin skull’ rule, which provides an exception to the ‘remoteness of damage’ rule. That holds that if some type of injury is foreseeable, but because of the particular predisposition of the victim the harm is greater than expected, the defendant is liable for the full extent of the injury rather than just the foreseeable portion.

Issues of actionability also arise at the level of basic moral philosophy. For example, when a wronged person seeks compensation from a wrongdoer, does her claim presuppose that she is made worse off by the wrong? Moreover, is the magnitude of compensation owed by the defendant determined by the extent to which the victim is made worse off? If we answer these questions affirmatively, a further question arises: by what standard of wellbeing do we resolve these issues? After all, competing conceptions of wellbeing will identify and measure compensation differently, and it is not obvious what view of wellbeing is correct. Is a person whose spouse is secretly unfaithful or whose publisher fails, against her wishes, to posthumously publish her life’s work rendered worse off by these events? A *Hedonist* would deny this, insisting that no event can make someone’s life go worse unless it affects her conscious experience by causing pain or unhappiness. A *Preference Theorist* would disagree, holding instead that one can be made worse off if one’s preferences are frustrated, irrespective of whether they affect one’s mental states. On this view, the person

whose spouse is unfaithful is made worse off even if she never discovers it.¹ Such theorists might agree that the spouse has been wronged, but disagree about how wellbeing is affected by the wrong.

As these brief remarks suggest, wellbeing is a complex and contentious philosophical concept. But the contested nature of wellbeing is not the only source of difficulty with respect to moral actionability. Even if we have established the correct conception of wellbeing – whatever that may be – it does not follow that rules regulating compensation should always invoke it. Perhaps respect for the victim’s autonomy requires that we defer to her conception of wellbeing, even if it is false. On the other hand, if the victim demands more compensation than is warranted by the correct conception of wellbeing, can we expect the defendant to bear extra burdens because of the victim’s mistake? More generally, what system of compensation is most justifiable to those subject to its authority, both claimants and defendants? This is the issue of actionability I will focus on here. It should be clear that actionability encompasses more than this, but this issue has been little discussed in the literature and is therefore worth exploring.

The relationship between the law and individuals’ conceptions of wellbeing is important for the philosophy of private law in which a theory of remedies is a central component, but their significance extends to the realm of political philosophy. Private law principles governing compensation may not allocate goods on the scale of redistributive mechanisms in the public sphere, but liability still allocates resources on a more local scale. Its operation is also coercive and is the site of deep disagreements about wellbeing. These parallels suggest that the correct principles in the realm of private law compensation may have implications for political justice, and *vice versa*.

¹ For discussion of competing theories of wellbeing, see Derek Parfit, *Reasons and Persons*, Oxford: Clarendon Press, (1984), appendix I and Thomas Scanlon, *What We Owe to Each Other*, Harvard: Cambridge, MASS, (2000), pp. 113 – 126.

In the political context, it is controversial whether the state should force or promote the correct conception of wellbeing amongst citizens. There is an important debate in political philosophy between perfectionists and anti-perfectionists. It is difficult to know precisely to which positions these terms refer. Broadly speaking, anti-perfectionists hold that the state cannot, in exercising its political power, offer as a justificatory reason any consideration grounded in a comprehensive set of moral, philosophical or religious values (or conceptions of the good). This includes a liberal comprehensive conception that promotes the value of autonomy. The anti-perfectionist constraint thus disables certain reasons from being legitimately offered as a justification for state action.

Liberal perfectionists deny this.² They argue that the state can justify its actions with reference to a comprehensive conception of the good based on liberal values. Liberal perfectionists may favour similar policies to anti-perfectionists, but they do so in order to promote valuable autonomous actions.³ So many liberals claim that, in its exercise of coercive power, the state ought not to promote any conception of the good while others claim that the state can refer only to liberal values, such as the need to foster autonomous choice amongst citizens, to justify state action. Although I will draw a number of points from this debate, I will avoid the terminology of perfectionism and anti-perfectionism. This is for two reasons. First, the terminology deployed in the debate is not always consistent, so it will be clearer to define the views considered in this chapter by stipulation. Second, *The Respect View* discussed later, which is the closest correlate of liberal political views, may well be favoured by both liberal perfectionists and anti-perfectionists. If that is true, it is unnecessary to attend to the finer disagreements between these two camps.

² See Joseph Raz, *The Morality of Freedom*, Oxford: Clarendon Press, (1988) for a defence of liberal perfectionism. For a more general account of perfectionism, see Thomas Hurka, *Perfectionism*, Oxford University Press, USA, (1996).

³ For a recent defence of anti-perfectionism against liberal perfectionism, see Jonathan Quong, *Liberalism Without Perfection*, Oxford: Oxford University Press, (2010).

Let us take a brief look at some of the reasons given in favour of anti-perfectionism to assess its appeal (some of these reasons also apply, in a modified form, to liberal perfectionism). If the state refrains from promoting a conception of the good, this is conducive to political stability as it ensures that citizens will not withdraw their support for the state on the grounds that it has used its power to further some conception that they reject.⁴ But the appeal of anti-perfectionism runs deeper than its role in securing this minimal form of stability. This is encapsulated by John Rawls' claim that an 'overlapping consensus' of reasonable conceptions of the good is not a mere *modus vivendi*, but a principled commitment to a political conception of justice that can be endorsed by all citizens.⁵ The notion of endorsement is hard to spell out, but it has intuitive force. It prevents citizens from being alienated from the state in its exercise of political power, and is closely connected to the problem of legitimacy. One constraint on the legitimate exercise of state power might be that such actions must be justified by political principles that can be endorsed consistently with various conceptions of the good. There is a great deal of literature that criticises and develops these ideas, which there is no hope of summarising here, but perhaps this outline illustrates the attractiveness of anti-perfectionism to many political and moral philosophers.

To my knowledge, these issues have not been discussed in the context of wrongful injuries in private law. Given the parallels mentioned above, it would be natural for liberals of both sorts to recognise the importance of a person's autonomous capacity to formulate their own views about wellbeing, which might influence the rules regulating compensation and actionability. But we must be wary of potential differences that undermine a simple parallel argument. Either way, a number of important issues turn on the outcome of this discussion. Liberals in the political context will wish to know the scope of their principles, and whether they can be extended to the arena of private law compensation. Scholars of

⁴ See John Rawls, *Political Liberalism*, New York: Columbia University Press, (1993) at p. 148.

⁵ *Ibid.*

private law will be interested in whether liberal concerns affect two central questions about compensatory damages: whether the loss suffered by the claimant is actionable and if so how much compensation ought to be paid.

Answering the latter questions is especially pertinent since these issues are almost non-existent in standard legal discussions of actionable damage. But they have implications for case law, even if they have not been clearly articulated by the judges involved. In *McFarlane v Tayside Health Board* [2000] 2 AC 59, for example, the House of Lords ruled that where a doctor's negligence in performing a sterilisation procedure resulted in the claimant having a baby that she would not otherwise have had, the claimant could not sue for the cost of bringing up the child. This decision was reached, in part, because the judges believed that it would be offensive for the law not to consider the birth of a healthy child a benefit to the parents that at least offset the financial burden of upbringing.⁶ If the liberal views are correct, it is crucial whether the child was a benefit to the parents on their own conception of wellbeing. On the facts of the case, Lord Slynn noted that despite their claim for compensation the respondents loved and cared for the child as an integral member of the family. There is a further difficulty here, as even if the parents accepted that the child was a great benefit once she was born, they may have had different views prior to the birth (which seems likely, given the decision for the husband to undergo a vasectomy). If so, at what point in time should the parents' views be taken into account? But besides this, there is a more general question: in cases of failed sterilisation where the parents consistently lack the belief that having a child improves their lives overall, is it acceptable for the court to impose on them a view about wellbeing that they do not endorse?

There is a complication here that is worth making explicit. There might be a distinction between how well one's life goes and what one has reason to value.⁷ Consider the

⁶ See Lord Steyn at p. 82 and Lord Millet at p. 111.

⁷ Derek Parfit draws on a similar distinction in his discussion of a case in which a young girl has a child early in life and is unable to give it a good start. Parfit argues that it would have been better if

couple who have a child a year earlier than planned. Once the child is born, the parents have special reasons to value that particular child. They might recoil from the idea of swapping their life with the child for the alternative in which they have a different child a year later. Does this mean that their lives have gone better by having *this* child than they otherwise would? Not necessarily. They can assert that their plans were frustrated and they missed out on all the things they were hoping to do that year. Assuming the child they would have had would have been just as fulfilling for them, it seems to follow that their lives would have gone better if their plans were successful. It is just that they have special reasons to value their life with the child they have, which means they value it even over an alternative in which their lives would have gone better.

Implicit in this distinction is the thought that we have special reasons to value *particular* objects, relationships, experiences etc. It might be objected that the reasons we have to value things are just constituents of how well our lives go. The parents' lives are better with the unplanned child *because* they have special reasons to value it. But although it is plausible that having special reasons to value things contributes to one's wellbeing, it must be remembered that the parents would also have special reasons to value the child they would otherwise have had – it is just that those reasons would have appeared at their preferred time. Similarly, if one person's child is more burdensome than another's, that may be one way in which the first person is worse off than the second, but it remains true that the first person has special reasons to value her life with her own child over the other's.

This raises the following question: in failed sterilisation cases, should the courts base claims for compensation on what people have reason to value or how well their lives go? It is hard to see how claims could be based on the former option. The special reasons we have to value our own children are very strong, and they would almost always rule out the

she had waited several years before having a child, even though the fact that she loves her child is enough to block the claim that she is irrational if she does not regret it. See Derek Parfit, *Reasons and Persons*, Oxford: Oxford University Press, (1986), pp. 360 – 1.

possibility of compensation, even when the parents' lives have become materially much worse. Moreover, the fact that the parents are worse off overall does not imply that they ought to value an alternative life over the one they have with their child. And *this* is presumably the 'offensive' implication that the judges in *McFarlane* were keen to avoid. In any case, to keep matters simple, let us focus only on claims motivated by a person's being made worse off, regardless of how this affects the reasons they have to value their current life over some alternative. Let us imagine that this was the nature of the claim in *McFarlane*.

If this stipulation is right, the decision was made because the parents were better off by having the child than they would otherwise be. One might doubt this decision on the ground that the judges' conclusions about the wellbeing of the claimants were wrong (a child is a whinging bundle of demands, one might think. How could *that* make one's life go better?) But this is not the kind of criticism I am interested in here. The liberal views are grounded in respect for the autonomy of individuals, rather than blanket skepticism about whether there *is* any correct view of wellbeing. We can therefore make progress on these issues by investigating hypothetical cases in which there is a conflict between the correct account of wellbeing and mistaken accounts held by the claimant or defendant.

There are four types of cases worth considering: (1) those in which the wrongdoer correctly believes that the victim is not made worse off (or is benefitted) but the victim mistakenly believes that she is harmed; (2) those in which the wrongdoer correctly believes that the victim is not made better off (or is harmed) but the victim mistakenly believes that she is benefitted; (3) those in which the victim is not made worse off (or is benefitted) but *both* the victim and the wrongdoer mistakenly believe that the victim is harmed; and (4) those in which the victim is not made better off (or is harmed) but *both* the victim and the wrongdoer mistakenly believe that the victim is benefitted. There are two central differences between cases of types (1) and (2), and cases of types (3) and (4). First, (1) and (2) involve a conflict between the conceptions of wellbeing held by the victim and wrongdoer, whereas in (3) and (4), both parties agree with each other's views on wellbeing. Second, in (1) and (2),

one party holds the correct conception of wellbeing whilst the other holds a false one, and in (3) and (4) both parties hold a false conception of wellbeing.

Here are three views that I will canvass in this chapter. *The Objective View* holds that compensation should always be determined by the correct conception of wellbeing, with no allowance for mistakes on either side. In (1), if the victim is benefitted, she is not entitled to compensation even though she believes she is harmed. In (2), if the victim is harmed, she is entitled to compensation even though she believes she is benefitted. The agreement between the claimant and defendant in (3) and (4) will not affect matters either. In (3), if the victim is benefitted, she is still not entitled to compensation even though *both* she and the wrongdoer believe she is harmed. And in (4), if the victim is harmed, she is still entitled to compensation even though *both* she and the wrongdoer believe she is benefitted. In all cases the correct view of wellbeing determines whether the victim ought to be compensated.

The Respect View places a premium on the victim's sincerely held beliefs about her wellbeing, even if they are false. This view holds that compensation should be paid in (1) but not in (2). In (1), even though the victim is mistaken, she can make a good faith claim that she is worse off. In (2), however, when she is harmed but believes she is benefitted, she cannot claim in good faith and so is not entitled to compensation. In (3) and (4), *The Respect View* has similar implications. In (3) compensation should be paid and in (4) it should not. In both types of case the victim and wrongdoer agree about wellbeing, even if they are both mistaken, so it would be wrong for the law to impose the correct conception on them.

So far, I have identified the implications of competing views about whether compensation ought to be paid. A second issue is *how much* compensation ought to be paid. *The Objective View* holds that compensation should be calculated according to the actual degree of harm done. So the correct conception of wellbeing determines both whether compensation is due and how it should be calculated. *The Respect View* holds that compensation should be quantified according to the victim's sincerely held view of wellbeing, an answer that is also consonant with its position on the first question.

I argue that both *The Objective View* and *The Respect View* are false. Instead, we should accept *The Mixed View*, which makes four central claims. First, in (1) and (2), there are reasons, which are sometimes decisive, to pay compensation. Where the victim is benefitted but falsely believes she is harmed, compensation should sometimes be paid to respect her autonomy. Where the victim is harmed but falsely believes she is benefitted, compensation should sometimes be paid to compensate for her harm. Second, in (3) and (4), there are reasons, which are sometimes decisive, to pay compensation. The judgements in (1) and (2) equally apply when the victim and injurer share the same mistaken conception of wellbeing. The fact that a victim sincerely believes she is harmed gives us some reason to compensate her which we should respond to even if she is not harmed. These reasons are not always decisive. If enforcing the wrongdoer's liability would make him too badly off, for example, this harm would outweigh respect for the victim's beliefs. Thirdly, wherever there is real harm done, damages should be calculated according to the correct conception of wellbeing, but where the victim falsely believes that harm is done, damages should be limited to a reasonable threshold. This claim means that respecting the victim's beliefs does not require that we *quantify* compensation according to those beliefs. Even if the reasons we have to respect autonomy are generally decisive when deciding whether she has a claim, they need not also govern the value of this claim. And finally, *The Mixed View* could be applied to all sincerely held false beliefs or it could be restricted to false beliefs that are reasonable in some sense. I will suggest, though only tentatively, that *The Mixed View* should apply to all sincerely held beliefs about wellbeing.

Here is a quick map of the argument. In section 1 I will briefly elucidate some tenets of liberal thought, with particular reference to Ronald Dworkin's work, to provide some clarification and contextualisation. In section 2 I will defend *The Respect View* against *The Objective View* with regard to (1) and (3)-type cases, although I will also argue that damages paid to respect a victim's autonomy should be limited. In sections 3 – 4 I will defend *The Objective View* against *The Respect View* with regard to (2) and (4)-type cases. If these arguments are successful, their combination will justify *The Mixed View*. In the conclusion I

will address the criticism that *The Mixed View* is an arbitrary compromise in its attempt to accommodate elements of both other views. I will seek to show that this perception is misleading: *The Mixed View* emerges from a principled assessment of the cases as a defensible, coherent and victim-centred approach.

1. Compensation and Dworkinian Liberalism

One attractive theory relating to compensation in the context of distributive justice is developed by Ronald Dworkin. This view offers a metric for judging whether one individual is worse off than another, which is a necessary (though not sufficient) condition for a claim to compensation under distributive justice.⁸ This metric consists of an ‘envy’ test designed to ascertain whether impersonal resources are distributed equally or unequally between individuals. The envy test is satisfied if no one prefers anyone else’s set of resources to her own.

The envy test has a number of functions within Dworkin’s broader political philosophy. For example, it provides an alternative to equality of welfare for interpersonal comparisons that is not beset by the problem of expensive tastes. One taste, A, is more expensive than another, B, if A requires consumption of more resources than B to reach the same level of welfare. This may be true, Dworkin memorably observed, of the person with a penchant for Plover’s eggs and pre-phylloxera Claret, but such a person should not be compensated for the lower relative welfare produced by her expensive tastes. Dworkin’s metric yields the plausible result in these cases, since the person with an expensive taste for rare claret does not typically envy the person who can achieve the same level of satisfaction from lemonade.

⁸ Ronald Dworkin, *Sovereign Virtue: The Theory and Practice of Equality*, Cambridge, MASS: Harvard University Press, (2000) at pp. 67 – 68 and 134 – 62.

For our purposes, however, its merits as an alternative to other egalitarian views need not detain us. For the present discussion, the important feature of this view is its treatment of autonomous choice. Unlike John Rawls' list of primary goods or the capabilities approach outlined by Amartya Sen and Martha Nussbaum, Dworkin's envy test does not prescribe any list of goods that are presumed to be in the interests of everyone. Instead, it allows each individual to bring her own conception of wellbeing, however idiosyncratic or erroneous, to bear on the question of whether she is worse off than another. Dworkin's view "allows us to cite, as disadvantages and handicaps, only what we treat in the same way in our own ethical life".⁹ Although this view is by no means universally accepted amongst liberals and egalitarians,¹⁰ it seems to respect individual autonomy uncompromisingly in its refusal to list *any* specific goods that are presumed to be either intrinsically good for everyone or instrumental in pursuing anyone's good.

The test proposed by Dworkin is comparative.¹¹ It judges whether one person is disadvantaged relative to another. This is the relevant test for the purpose of distributive justice, the argument goes, as it is hardly warranted to compensate someone who believes he is better off *compared* to everyone else, but still badly off *non-comparatively* because he does not live in a society that values and subsidises his tastes. One who firmly believes that his life can only flourish if he hears regular performances of Rachmaninov's *oeuvre* may

⁹ Dworkin, *Sovereign Virtue*, p. 294.

¹⁰ Two notable proponents are Mathew Clayton and Andrew Williams. See "The Resources of Liberal Equality", *Imprints*, 5, (2000), 63 – 84 and "Equality for the Ambitious", *Philosophical Quarterly*, 53, (2002), 377 – 89 respectively. For a critique of Dworkin from a luck egalitarian perspective, see G. A. Cohen, *On The Currency of Egalitarian Justice and Other Essays in Political Philosophy*, Princeton: Princeton University Press, (2011) and "Expensive Taste Rides Again" in *Dworkin and his Critics*, Ed., Justine Burley, Blackwell Publishing, (2004).

¹¹ In this paragraph I follow Mathew Clayton's characterisation of the envy test. See "The Resources of Liberal Equality", pp. 77 – 8.

consider himself better off than the philistines who dismiss Rachmaninov, but not very well off overall because of the paucity of such performances available to him. Whatever sympathy we may have for such a person, he has no claim against other citizens to be compensated for the low welfare attributable to his unusual circumstances.

The distinction drawn before between how well someone fares and what she has reason to value might threaten the envy test. Whether one individual envies another depends largely on what she has reason to value, but it is arguable that distributive justice (or resource allocation) should depend on how well off she is. Consider a society divided by a caste system. Birth into a particular caste determines one's opportunities, resources and social treatment, but also what set of cultural values and practices one inherits. A member of the lowest caste might be born into poverty and servitude, but refuse to swap places with a member of a higher caste. This choice is not inconceivable or even irrational; the individual might have particular reasons to value her own cultural heritage, even if it is also the source of material deprivation. It is plausible that this individual is worse off than those in higher castes even though she does not satisfy the envy test. Moreover, she has a claim to be compensated as a matter of distributive justice. This suggests that justice should depend on how well off a person is rather than what they have reason to value, or whether they envy others in the Dworkinian sense. I will not attempt to settle this issue here. To avoid complications, I will restrict my focus to judgements about how well off the victim is, rather than what she has reason to value or whether she satisfies the envy test.

The interpersonal comparativism exemplified by the envy test may be appropriate for claims in distributive justice, but it will not do for wrongfully caused injuries. The primary ground for private law claims is not inequality. But we can rectify this, holding instead that the victim must be worse off than *she* might have been, according to her own conception of wellbeing.¹² With this modification, our test determines whether a person

¹² I assume here that the appropriate object of comparison with the victim's actual state is a relevant counterfactual state. In making this assumption I am implicitly denying non-comparative accounts of

believes herself to be worse off compared to how well off she would be in a relevant counterfactual state, and is thus a suitable metric to apply to cases of wrongfully caused injury.

Thus, we arrive at *The Respect View*. This view holds that where the victim believes, in good faith, that she is made worse off by the wrong than she would otherwise be, we have reasons to recognise a moral duty on behalf of the wrongdoer to offer compensation, and the amount of compensation owed should be quantified by the victim's conception of wellbeing.

Before evaluating this view, two caveats and clarifications must be noted. First, although it seems to capture the intuition behind the envy test in the context of wrongfully caused harm, I do not mean to attribute *The Respect View* to Dworkin or anyone who defends his account of interpersonal comparison, if only because this account has never been explicitly applied to the kind of cases I will discuss. Secondly, remember that *The Respect View* purports to be decisive only if it is established that the defendant has committed a wrong and has no recourse to a defence. It is not a complete account of liability. It only addresses two elements of liability: whether the victim has suffered the kind of consequences that merit compensation and which conception of wellbeing should feature in the quantification of damages

2. Defending *The Respect View* against *The Objective View*

The Respect View has plausible implications in many cases. In particular, it captures something important when applied to (1) and (3)-type cases, where the victim is not made worse off (or is benefitted) but she mistakenly believes that she is harmed. Respecting the

harm or disadvantage, as well as other (non-counterfactual) comparative accounts such as temporal comparativism. I attempted to justify these assumptions in chapter 3.

autonomy of the victim often demands that she be compensated even though she has been benefitted.

There are many different cases that fit these schemas, and how one will describe these cases obviously depends on one's views about wellbeing. If we agree with the judges in *Mcfarlane* that a healthy child is a benefit, then in failed sterilisation cases where the parents reject this judgement, the victims have false beliefs about their wellbeing. We can also construct cases involving more general disagreements about the nature of wellbeing. Here is a hypothetical example based on a more fundamental debate between philosophers who endorse hedonistic and non-hedonistic views:

The Reticent Philosopher: V spends her life writing a series of philosophical works. Although the works are brilliant, advancing knowledge in various ways, she decides not to publish them as she finds publicity excruciating. D negligently distributes the works against the author's wishes, which deepen the understanding of many who read them and thrusts her into the limelight. V, however, is a *Hedonist* about wellbeing and does not believe that the achievement improves her life. Instead, she finds her new life in the public eye unbearable.

Let us assume that *Hedonism* is false and V's philosophical achievements make her life go better. Moreover, these achievements are significant enough to outweigh any discomfort she feels about her fame. Of course, these are controversial assumptions about the nature of wellbeing, but this is not problematic. We are simply stipulating that these claims are true in order to ask what follows. We could easily reformulate the case on the presumption that some other view of wellbeing is correct.

Like the failed sterilisation examples, in *The Reticent Philosopher*, D benefits V by wronging her. V's hedonistic view is wrong: her life is better as a result of her achievements even if this is not reflected by any improvements in the quality of her mental states. Despite this, it is intuitive that V's choices should be respected in both cases, and indeed any other (1) and (3)-type cases, and we should recognise a moral duty owed by D to offer

compensation. D might object to this on the grounds that, in truth, he has benefitted her. He might argue that it is reasonable to expect victims to take responsibility for their own mistakes, and the burden for this mistake should not fall on him. Thus, even though he accepts responsibility for his wrongful conduct, which in principle involves payment of damages for any harm done, he owes no compensation as he has done no harm.

But we have two important reasons to compensate V nonetheless. First, the wrongdoer can be expected to bear some costs for a system of compensation that respects the autonomy of victims. The liberal intuition comes into its own in these cases. If compensation is measured according to a single legally sanctioned metric, with which many citizens fundamentally disagree, this policy could not be justified to those on whom it is imposed. It would be vulnerable to familiar criticisms of general illiberal policies.¹³ Many people find this idea very appealing, but as I said at the start, it is quite difficult to spell out its underlying rationale. It is therefore worth pausing to consider a couple of views that seek to explain its appeal.

One view, attributable to Jonathan Quong, can be called *The Moral Status View*. This makes two claims. First, the law's rejection of the victim's false conception of wellbeing treats her as if she lacks the capacity, in this context, to plan, revise and rationally pursue her own conception of the good. The second claim is that to treat her in this way is inconsistent with her moral (or political) status as a free and equal citizen.¹⁴ This explains the liberal intuition that V's sincerely held beliefs should not be discarded.

Both of these claims might be questioned. To take the second claim first, to know whether such treatment is inconsistent with a person's moral status, we need to know more

¹³ See, for example, Jon Quong on the charge of paternalism in *Liberalism Without Perfection*, Oxford: Oxford University Press, (2011), pp. 73 – 108. I will later argue that paternalism is not always problematic, although I agree that it is problematic in (1) and (3)-type cases, as it involves a violation of the victims choices.

¹⁴ See *ibid.* pp. 100 – 6.

about the nature of this status. The law regularly denies that certain individuals lack the capacity to rationally pursue a conception of the good when it restricts the civil liberties of the young or mentally disabled. It is not clear that in doing so it fails to respect their moral status. We do not, after all, consider the mentally disabled to be unequal in terms of moral status. Nevertheless, a defender of *The Moral Status View* could build a plausible argument that treating those who have false beliefs about their wellbeing as equivalent to the mentally disabled is inconsistent with their moral status, so let us focus on the first claim.

As a point of clarification, it is surely wrong for Quong to say that refusing to compensate victims according to their mistaken beliefs treats them as if they lack the capacity to pursue their own conception of wellbeing.¹⁵ This is nonsensical: avoiding their conception of wellbeing is precisely what these judgements aim at! The claim is rather that they treat victims as if they lack the capacity to form the *correct* conception of wellbeing.

This claim might be criticised in two ways. First, it might be questioned whether this treatment is deeply problematic. Individuals are still free to make their own mistake as long as they do not increase the burdens that others are expected to bear for them. It is not unreasonable, this objection holds, to judge that victims are unable to form correct beliefs about wellbeing when the burdens of others are affected by these beliefs. Second, it might be doubted whether in these cases victims are really treated as if they lack the capacity to pursue the correct conception of the good. Why, one might ask, should accurately judging that they have made a mistake imply this? It is a basic truth that judging that someone has got something wrong does not imply that they could not have got it right. If I notice that a student has made a mistake in an exam, this does not imply that they could not have reached the correct answer. It is therefore too strong to conclude that victims are treated as if they lack the *capacity* to pursue a valid conception of the good.

I will not explore this view any further here. Instead, let us briefly consider an alternative that seeks to explain the appeal of the intuition behind *The Respect View*. This

¹⁵ Ibid. at p. 101.

can be called *The Endorseability View*. It states that we have reasons to ensure that the set of laws to which citizens are subject are endorseable from their perspective. Joseph Raz summarises the attraction of this view: “While I am concerned to lead the life that I now believe to be the right one, I am even more concerned to be able to lead the life that conforms to my freely developed conception of the good as it may be from time to time. Hence the way to relate to others who do not share our conception of the good is to establish a scheme of cooperation, to which all could agree, and which would enable all to pursue their own conceptions of the good within fair terms of cooperation.”¹⁶

This might be beneficial in a number of ways. For those, like Raz, who believe that a political doctrine of liberal justice is embedded in a true comprehensive moral theory, the state facilitates the value of autonomously chosen actions by ensuring that its laws are endorseable. The appeal of endorseability is thus partly explained by its relation to a fundamental moral doctrine. As Raz explains, there is a double benefit on this view, as this fundamental doctrine is even consonant with the views of those who reject it. In fact, it is the correct doctrine *because* even those who reject it are implicitly committed to it. In satisfying this condition, it is the doctrine that best realises the value of allowing citizens to pursue their freely chosen ways of life.

If *The Endorseability View* is right, it also suggests that we have reason to respect perverse or wildly wrong beliefs as well as reasonable ones. It is important that citizens can endorse the entire set of laws that apply to them, rather than particular provisions or judgements. It is difficult to see how this could be achieved if the courts made it their business to make fine or controversial distinctions between false beliefs that are reasonable and those that are too abhorrent to merit respect. And of course, if such distinctions were made, those with unreasonable false beliefs could not endorse the law as a whole.

¹⁶ Joseph Raz, “Facing Diversity: The Case of Epistemic Abstinence”, *Philosophy and Public Affairs*, 19(1), (1990), 3 – 46, p.26

Since we lack the space to adjudicate between these views, it is probably best to remain ecumenical, although the above comments indicate that there are some serious problems with *The Moral Status View*. Both purport to explain the intuition that drives *The Respect View*, and support more general constraints designed to prevent state power being exercised for the furtherance of comprehensive conceptions of the good that can be reasonably rejected. Hopefully this brief sketch gives some explanation, however patchy, of the liberal underpinnings of *The Respect View*.

The second reason we have to recognise D's moral duty to offer compensation is that D cannot reasonably *complain* about bearing these burdens for the sake of respecting V's autonomy. He only bears this burden because of his own conduct, and given this conduct, it is inevitable that the extent of his liability is partly a matter of luck. Can he really complain when he might just as easily have harmed the victim according to the correct conception of wellbeing? He could have avoided running this risk by refraining from his negligent conduct (note that the avoidability argument is even stronger if his wrongful act is intentional, as it is easier to avoid intentional wrongs). His grounds for complaint are weaker given that he was in the best position to avoid his own liability.

We have assumed that *The Reticent Philosopher* is a type-(1) case. In other words, D does not hold a false view and knows that he has benefitted V. The considerations outlined above justify the rejection of his view for the sake of V's autonomy. D's grounds for complaint are further undermined if we turn *The Reticent Philosopher* into a type-(3) case in which the defendant shares V's false view of wellbeing and believes he has harmed her. It is not necessary that D's beliefs should also be false, but if they are this can only be another point against his complaint.

If the foregoing discussion is correct, *The Respect View* shows that a common assumption is false. That assumption is that it is necessary for a claim to compensation that the wrongdoer causes harm to the victim. One premise here is that the quantification of damages follows from the measure of harm. If it does not, this raises a quandary, as how is compensation quantified if not by measuring the amount of harm done? But if V has a valid

claim in *The Reticent Philosopher* or failed sterilisation cases, the assumption must be false: V is owed compensation even though she has not been harmed.¹⁷

3. Quantifying Compensation

Given that the victim has an actionable claim in (1) and (3)-type cases, how much should she be compensated? *The Respect View* holds that she is entitled to be compensated according to her own beliefs. In practice, of course, such a rule would often have to be displaced by an objective calculation because, in the absence of any convincing evidence, it would be too

¹⁷ Admittedly, this claim depends on which view of harm we adopt. Those who believe that harm consists in the frustration of preferences will hold that V has been harmed even though her beliefs are false. I reject this view of harm in chapter 3. I will note, however, that the *possibility* of cases such as *The Reticent Philosopher* provides a further reason to reject preference-based accounts of harm. It is surely possible that a person's false beliefs can make their preferences conflict with what is genuinely to their benefit. The preference-based account implausibly dismisses these cases as conceptually impossible.

Moreover, in outlining this reason why the assumption that it is necessary for compensation that D harms V is false, I do not mean to suggest that it is the only reason. Briefly, here is another. Suppose V is in an accident and is taken to hospital for a skin graft. While she is unconscious, the surgeon performing the procedure notices that her nose could be more attractively shaped and takes the liberty of fixing it (see Matthew Clayton's "The Case Against the Comprehensive Enrolment of Children", *The Journal of Political Philosophy*, 20(3), (2012), 353 – 364 at p. 357). In this case it is plausible that V ought to be compensated even if she believes (correctly, let us assume) that she has been benefitted by the procedure. V had a right that her consent be sought and the failure to do so grounds her claim. V's claim is based neither on harm she has suffered (since she has benefitted) nor on her sincere belief that she is harmed (since she does not hold this belief).

easy for claimants to falsely report their beliefs. But even in principle, we ought to reject this implication. This is for two reasons. The first reason is not particular to *The Respect View*, but sets a limit on actionable damage generally. We are often not permitted to enforce compensatory duties if this will place the defendant below a certain threshold of wellbeing. Suppose D steals a small amount from V, but fully compensating V will render D destitute. It may not be permissible to enforce D's liability to V because doing so will leave D too badly off.¹⁸

This threshold places a general limit on what liabilities we are permitted to enforce and the extent to which we may do so. A full discussion of this threshold would deviate too far from the evaluation of *The Respect View*, which is our present purpose, but it is worth mentioning one question that arises in this context. Suppose that when D steals from V, compensating her (according to the correct conception) requires paying £X to V. Imagine too that compensating V in *The Reticent Philosopher* (according to her false conception) also requires paying V £X. If we are permitted to enforce D's liability to pay £X in the first case, does it follow that we are permitted to do so in the second? Imagine that requiring D to pay £X in the first case would be *just about* acceptable. Having to give up the money would leave D badly off, but not so badly off that third parties are not permitted to force him to correct his wrong. As least, one might think, he is compensating for the harm he has inflicted. Is the same true of the second case? Are we permitted to force D to transfer the money to V, leaving him equally badly off, in order to compensate V according to her false beliefs even though D has in fact benefitted her? In other words, does the truth of the matter about whether D has benefitted or harmed V affect the threshold that limits enforceable liabilities?

It seems clear that it does. Consider:

¹⁸ See Victor Tadros, "Orwell's Battle with Britain: Vicarious Liability for Unjust Aggression" (forthcoming), p. 9.

Lightning Bolt: *Achilles* and *Narcissus* are struck by a lightning bolt negligently issued from a conjurer whilst trying to impress a group of children. *Achilles* receives an injury to his left heel, which paralyzes him, while *Narcissus* receives a shallow wound. *Narcissus*' wound leaves a faint scar, but as he is extraordinarily vain, he sincerely believes that he is harmed just as much as *Achilles* and demands an equal amount of compensation. The conjurer has enough to pay both victims, but as his coffers are running dry, doing so would leave him very badly off.

In *Lightning Bolt*, it is intuitive that *Narcissus* should not be granted the award he seeks, but *Achilles* is entitled to full compensation. This asymmetry must be based on the judgment that *Narcissus*' beliefs about his wellbeing are false. We are not willing to render the conjurer so badly off to respect

Narcissus' false beliefs simply because he is negligent, but we are willing to do so in response to the real harm inflicted on *Achilles*.

The next reason relates to the size of D's liability. When D wrongs V, it is foreseeable that V will be harmed. However, the extent of the harm may not be. If V has some condition that makes her harm greater than that which is foreseeable from D's *ex ante* perspective, it seems fair to require D to compensate her for the entire injury, including the unforeseeable portion. In law, this is known as the 'thin skull rule', which I mentioned in the introduction.

The Respect View might be seen as an extended version of the thin skull rule. Some victims have thin skulls, others have thin beliefs. Either way the defendant must take her victim as she finds him. But now return to *Achilles* and *Narcissus*. Even if D can compensate both of them without rendering himself very badly off, it is still implausible that *Narcissus* deserves as much compensation as *Achilles*. It is not just the effect of the payment on D's wellbeing that prevents us reaching this conclusion. If we treat *Narcissus*' affronted vanity (especially if we are dealing with legal as well as moral duties) as equivalent to *Achilles*' paralysis, this is an insult to *Achilles*. These considerations undermine *The Respect View*.

Even if we accept that *Narcissus* has an actionable claim, it is wrong to suggest that compensation should be quantified according to his false beliefs. Instead, the responses sketched above support the following claim. Where a victim suffers actual harm, compensation should be quantified according to the correct account of wellbeing, but when a victim does not suffer harm or is benefitted, any claim he possesses, grounded in respect for his autonomy, is limited to a reasonable threshold.

I will not discuss this threshold in any more detail. At present I am only arguing that it exists. However, we are still faced with the quandary I mentioned earlier. If the victim has an actionable claim grounded in respect for her autonomy, yet she is not entitled to the sum of damages she sincerely believes she deserves, how are her damages to be calculated? One solution to this problem would be to stipulate a threshold that is consistent between claimants. This would allow the victim's beliefs to play at least some role in the ideal quantification, but only up to a limit. Another solution would be to offer all victims a fixed sum to recognise and make allowance for their beliefs about the harm they have suffered. Tort law adopts a similar policy elsewhere, where it is difficult or undesirable to quantify harms, such as damages awarded for bereavement under S.1(a) of The Fatal Accidents Act.

As before, these conclusions are not affected if we imagine that the case of *Achilles* and *Narcissus* is type-(3) rather than type-(1). If evidence could be offered that the conjurer is also extraordinarily vain and falsely believes that *Narcissus* is as badly off as *Achilles*, such evidence could not be used to force him to pay *Narcissus* as much as *Achilles*. The central point is that it is unfair to force him to pay an enormous sum of compensation as a result of *Narcissus*' mistake, even if this mistake is also made by him.

4. Defending *The Objective View* against *The Respect View*

So far we have seen that *The Respect View* has plausible implications in cases where D wrongs V, D's wrong benefits V but V believes that she is worse off (or not better off) in

virtue of her false conception of wellbeing. In this type of case, V's good faith judgment is sufficient for her claim, subject to the limitations described above.

What about (2) and (4)-type cases where D wrongs V, D's wrong harms V but V believes herself to be better off (or not worse off) according to her false beliefs? *The Respect View* holds that V has no claim here. Is it correct to suggest that it is *necessary* that V believes she is made worse off before she is able to claim? It must be remembered that the question is whether D owes a moral duty to offer compensation to V, not whether she must accept. Certainly, neither D nor the court can *force* V to improve her wellbeing; it is up to V whether she wishes to seek recompense for D's wrong. Assuming this moral duty could be transferred to the legal context, it should also be recognised that we are therefore not merely interested in V's preferences as they are reported in court (the issue will only arise if V decides to seek damages) but as they are revealed in her behaviour. So assuming V seeks the maximum damages available to her, the question is this: if D could provide decisive evidence that he has not harmed her by her own standards, can D rely on this to escape paying damages to V?

Before testing its implications in a particular scenario, it is worth mentioning a few general points that reduce the sting of *The Respect View* in these cases. When is the intuition behind this view at its strongest? We can investigate this by taking Rawls' political liberalism as a parallel. For Rawls, the state's neutrality between reasonable conceptions of the good is an integral part of what he calls the publicity condition.¹⁹ The details of the publicity condition need not concern us. What are relevant are the circumstances in which it is seen as necessary. That is, the circumstances in which the neutrality constraint, which enables principles to be justifiable to reasonable citizens who endorse plural and conflicting conceptions of the good, must be met in order for political power to be exercised legitimately.

¹⁹ See John Rawls, *Political Liberalism*, New York: Columbia University Press, (1993), pp. 66 – 71.

These circumstances are characterised by four features that apply to the citizen's relationship with the basic structure of society. This relationship is *involuntary*, *coercive*, *lifelong* and *profoundly impactful* on "citizens' character and aims, the kinds of persons they are and aspire to be."²⁰ The publicity condition should in principle apply whenever these four features are present, and others have argued that liberal legitimacy ought to be extended to other domains beyond those specified by Rawls.²¹

Do they apply to cases of wrongful harm? Principles regulating compensation are involuntary and coercive for victims who lack a decent opportunity to avoid exposing themselves to harm. If they are unlucky enough to be wronged, they have no influence over the principles that will determine whether and how much they are compensated. This is one reason why *The Respect View* gives plausible results in cases where V demands compensation even though she is benefitted. But the same is not true of *defendants* in the kind of cases we are now considering. It is true that the law is coercive *once* a defendant has incurred liability: the claimant is given a power exercisable against the defendant to co-opt the court in enforcing the defendant's corrective duties. But defendants generally have decent opportunities to avoid committing wrongs and therefore to avoid being subject to the coercive authority of the court. Although committing a wrong cannot be taken as a sign that a defendant voluntarily accepts the legal consequences, the opportunity he has to avoid putting himself in this position substantially reduces the liberal worry about the justification

²⁰ Ibid. at p. 68.

²¹ See, for example, Mathew Clayton, *Justice and Legitimacy in Upbringing*, Oxford: Oxford University Press, (2006), especially pp. 93 – 124 (applied to the upbringing of children) and Joshua Cohen, "The Economic Basis of Deliberative Democracy", *Social Philosophy and Policy*, 6(2), (1989), 25 – 50 (applied the democratic defence of socialism).

of laws that coerce those to whom they apply.²² These considerations do not refute *The Respect View*, but they diminish its urgency and appeal.

Let us now consider some cases. As before, these will depend on particular views about wellbeing. For example, if you rightly think that Mother Teresa was deeply misguided to believe that physical suffering is a benefit because it brings the victim closer to identifying with Jesus Christ, injuring a devout follower of Mother Teresa will be a case in which the victim is harmed but mistakenly believes she is benefitted. If you are an atheist, you might agree that if a person hits her head, has a religious epiphany and decides to become an evangelical missionary, her life goes worse even though she falsely believes it goes better.

In the cases we use to test the plausibility of *The Respect View* the victims' false beliefs should have non-negligible implications for their wellbeing. This imposes at least two conditions on the construction of hypothetical cases. First, we must be wary of benefits that flow from false beliefs. Even if having true beliefs, *ceteris paribus*, makes one's life go better than having false beliefs, it is possible that this disadvantage is outweighed by other benefits flowing from those beliefs. This may be true of someone who derives enormous comfort and pleasure from her religious beliefs, and is generally uninterested in truth and intellectual inquiry. To circumvent these complications, we must only consider cases where the victim's false beliefs are bad for them all things considered.

There is one other possibility about which we must be aware. *The Respect View* requires that we do not impose on the victim a judgment about her wellbeing, and this

²² It must be remembered, however, that this consideration applies more strongly to intentional than unintentional harm: forming bad intentions is usually more avoidable than taking negligent risks. In turn, unintentional wrongs are usually more avoidable than risks created by permissible activities. Nevertheless, it is also generally true that negligent actors have better opportunities to avoid the harm they cause than their victims (and the same is sometimes true of agents who cause harm in the course of permissible conduct).

requirement is violated regardless of how severely her false beliefs affect her wellbeing. If we only examine cases in which her false beliefs do not have severe implications, *The Respect View* may give plausible results. But this may not be because it is true but rather because it is partially coextensive with some other principle. For instance, there may be a sufficientarian logic operating here. Perhaps, as long as one's false beliefs do not have too great a negative impact on one's welfare, autonomous choice must be respected. Autonomy only has priority once a threshold of objective wellbeing is reached.²³

With these caveats noted, consider the following:

The Shallow Philosopher: V is desperate for fame. The easiest avenues for seeking fame are closed off to her as she is terrible at everything but philosophy. She spends her life writing a series of philosophical works which she hopes to publish with a reputable publisher. The works are brilliant, advancing knowledge on a variety of philosophical problems. But V's publisher, seeking wider circulation, wrongfully turns the works into a rhetorical and hysterical pamphlet. The pamphlet is successful: it is widely bought and distributed, dumbing down the opinions of readers all over the world. V becomes internationally famous and enjoys her newfound success. As she is a *Hedonist* about wellbeing she believes that she is made better off by the distribution of the pamphlet than she would have been had the original works been distributed.

Again, we must rely on a number of stipulations here. We must assume that *Hedonism* is false. D makes V's life go significantly worse by ruining her great

²³ For an argument that there should be at least some restrictions on liabilities incurred by individual choices, grounded in duties to live with dignity, to respect the value of rational agency and to protect valuable social relations, see Andrew Williams, "How Gifts and Gambles Preserve Justice", *Economics and Philosophy*, 29(1), (2013), 65 – 85.

achievements even though this does not negatively affect her mental states. Moreover, we must assume, not implausibly, that the transformation of a great work of philosophy into a flamboyant pamphlet considerably diminishes its value. *The Respect View* holds that D does not owe any compensation to V as V cannot make a good faith claim that she is worse off. This is an implausible implication. Furthermore, this judgement is not restricted to *The Shallow Philosopher*. It is also implausible that the injurer owes no duty to offer compensation in the Mother Teresa and evangelical epiphany cases noted above. This conclusion is, I suggest, generalisable across (2) and (4)-type cases. There are a number of reasons why we should reject *The Respect View* in these cases.

As before, the defendant could easily have avoided being burdened by the mistakes of others. It is only because of his wrongful conduct that he is subject to coercive rules that force him to compensate his victim. Unlike duties in distributive justice, his duty is incurred as a result of his positive, avoidable and wrongful conduct. This diminishes whatever complaint he might otherwise have against compensating a victim who mistakenly does not believe that she is harmed.

Secondly, we have paternalistic reasons to compensate V. V's beliefs may eventually change, perhaps as a result of the award, so that they align with the correct conception of wellbeing, and this gives us a reason to compensate her. We have these reasons even if her beliefs will not change, as long as the award makes her better off regardless. This may seem initially unappealing to those liberals who deeply object to paternalism in political theory, but there are several fundamental differences that make these paternalistic reasons less problematic than paternalism in the political sphere. The most troubling examples of paternalism involve coercing a person in contradiction to her choices. In *The Reticent Philosopher* and failed sterilisation cases, recognising that D has benefitted V would require violating her choices by flatly refusing her request for compensation, and so autonomy-respecting reasons override paternalistic ones. In *The Shallow Philosopher* and the Mother Teresa case, the victim is not coerced and her choices are not violated. The

question is just whether D owes a moral duty to offer compensation. After all, it is at the victim's discretion whether she seeks to enforce D's duty.

Paternalism is also worrying if one attempts to influence the behaviour of another for their own benefit by making their preferred choices harder. For example, the state may tax undesirable activities, making them difficult to pursue, even if it does not prohibit them. This charge of paternalism is also inapplicable here. By offering the victim compensation, the defendant does not make it more difficult for her to pursue her false conception of wellbeing; it merely makes it easier to pursue the correct one. It gives her an *opportunity* to improve her wellbeing which she does not have to take (although given that money is an instrumental good, she may use the award to improve her wellbeing in some other respect. I will return to this possibility below).²⁴

²⁴ There is disagreement about how best to define paternalism. Violating a person's choices for their own benefit is paternalistic on any almost any definition. Perhaps the most restrictive definition is that a paternalistic act limits a person's liberty. Writers who accept the liberty-limiting definition include Gerald Dworkin and Richard Arneson. See "Paternalism," *The Monist*, 56, (1972), 64–84 and "Mill versus Paternalism", *Ethics*, 90, (1980), 470–89.

Jonathan Quong rejects this definition because it is too restrictive. Encouraging someone to make the right choice might be paternalistic without limiting liberty (see Jonathan Quong, *Liberalism Without Perfection*, Oxford: Oxford University Press, 2010). Instead he adopts 'the judgmental definition', according to which an act is paternalistic if it seeks to improve someone's situation and is motivated by a negative judgement about that person's ability to make a competent decision for herself. As I claimed before, judging that a person has a false belief does not entail the further judgement that a person lacks the competence to form true beliefs. So perhaps Quong's definition is too restrictive after all: encouraging someone to make the right choice because you think they might get it wrong is paternalistic, but it does not involve judging that they lack the competence to get it right (perhaps the encouragement simply aims to reduce the risk). An act motivated by the judgement that the beneficiary is not omniscient is hardly paternalistic.

Perhaps a victim will use her compensation to make herself even worse off. Imagine that the person in *The Shallow Philosopher* is so irreversibly steeped in her hedonism that she uses the money to buy an experience machine, making her life go even worse. In principle, if a defendant knows or has good evidence that this is the case, this will undermine the paternalistic rationale for offering compensation. In practice, it is unlikely that a defendant would have such knowledge, but we can accommodate this possibility by making the duty conditional. The duty is engaged only if there is no good evidence that the victim will use the resources to make herself worse off.

5. The Bad Faith Objection

There is an important objection to the argument of section 4. If V can claim compensation even if she believes herself to be better off, then she is claiming in bad faith.²⁵ Presumably when one claims compensation, one implicitly agrees that one *has* been harmed, otherwise what is the purpose of the award? In *The Shallow Philosopher*, there is something duplicitous about V's claim. She seeks compensation from D even though she has not been harmed according to her (albeit incorrect) conception of wellbeing. Isn't *The Respect View* right to insist that any successful claim must be made in good faith?

First of all, the bad faith objection is not an objection to recognising a moral duty on behalf of the wrongdoer to offer compensation; it is only an objection to the victim's

Whatever the precise definition of paternalism, offering compensation to someone who does not believe that they are harmed seems to be paternalistic. My claim, therefore, is not that the implications of *The Objective View* in this case are not paternalistic after all, but rather that they are acceptably paternalistic.

²⁵ For Dworkin's discussion of the good faith constraint in his political theory, see *Sovereign Virtue* at pp. 148 – 61.

standing to enforce that duty. Even if the objection has force, therefore, a number of other responses to the duty besides enforcement by the victim would be available. The court could make a public declaration that the defendant owes such a duty and ought to fulfil it. There is case law to suggest that civil courts currently have jurisdiction to make similar declarations as a way of vindicating the rights of a victim, even if no compensation is transferred.²⁶ Alternatively, the defendant's duty could be enforced by those who do not lack standing to do so, such as the victim's family members or descendants. Having narrowed the scope of the objection, let us now see whether it has any force.

In some cases, victims may not be so firm in their beliefs. V might recognise that she is very badly off if her views about *Hedonism* are false. She would not be claiming in bad faith if she sought compensation as a form of insurance, to improve her wellbeing in other areas, just in case she had the wrong belief. This is not necessarily a duplicitous expression of bad faith; it might express of a degree of uncertainty in her convictions about wellbeing. Furthermore, it does not follow from this that she does not *sincerely* believe that she is not harmed. This would mischaracterise the nature of her belief. She does not believe that she is harmed by the ruination of her novels, even if this belief is not absolute. Absolute certainty, or even substantial certainty, is not a necessary condition of belief. Admittedly, this doxastic thesis on the conditions of belief requires more of a defence than I can offer here, although there are plenty of examples where it is intuitive that one can believe without absolute confidence (think of recollection: I can believe something about a past event even if my memory is sufficiently hazy that I lack confidence in this belief).

Moreover, as I argued in section 2, we should not assume that all compensation claims assume that the victim has in fact been harmed. Nevertheless, in cases like *The Shallow Philosopher*, at least the victim can sincerely claim that she is worse off, even if this claim is false. But there are other reasons why compensation is not always grounded in actual harm. For example, in the English law of damages, a claimant's compensation is not

²⁶ See *Ashley and another v Chief Constable of Sussex Police* [2008] UKHL 25.

always reduced to deduct benefits flowing from the wrong. But this would be illogical if compensation is uniformly a response to harm. Charitable gifts to victims are causally dependent on the wrong having occurred, and in this respect, the defendant benefits the victim by wronging her. But such benefits are not deducted from the claimant's compensation.²⁷ This may be for a number of reasons. Perhaps it is unfair if the wrongdoer benefits from the charity of others. Also, if such benefits are deducted, this may disincentivise charitable gifts to victims. Either way, the victim receives compensation for aspects of the harm that have already been compensated.

Rules like this should encourage us to be clearer about the content of the good faith requirement. The objection assumes that V must claim in good faith that all of her compensation responds to harm that she claims has been inflicted on her. This is not true of charitable gifts, since the victim can accept that she is doubly benefited (once by the gift and again by the compensation from which the gift is non-deductible) without the kind of bad faith that would undermine her claim. Given this, we should be less dogmatically opposed to compensation that does not always respond to harms recognised by the victim.

In any case, it is not clear that good faith is generally necessary for valid compensation claims. First note that it is not problematic to condemn interfering with someone as wrongful or impermissible even if the victim believes he benefits from it. Suppose that V is a sado-masochist and enjoys others inflicting pain on him. V wants D to inflict pain and would consent to this if D asked. On the basis of this hypothetical consent, D wounds him. Clearly it was not permissible for D to do this despite the fact that V believes himself to be benefitted. Equally, it is not permissible for doctors to touch or perform surgery on patients without their consent, even if the patients believe such actions are to their benefit and would give consent if asked. And finally, it is not permissible for D to

²⁷ See *Parry v Cleaver* [1970] AC 1 (HL). Although note that if the claimant is given gratuitous care by family she can sue for normal cost but holds recovery in trust for the carer, see *Hunt v Severs* [1994] 2 AC 350 (HL).

distribute the work in *The Shallow Philosopher* even though V does not believe that it makes her worse off.

Moreover, the *defendant* in type-(2) cases does not subscribe to an erroneous conception of wellbeing. From his perspective, he will have failed to fulfil his corrective duties if he does not offer compensation to his victim. There is also a great deal of bad faith on behalf of a defendant who knows he has harmed someone, has been sued by his victim, but refuses to pay damages by appealing to the victim's conception of wellbeing that he knows to be mistaken. It is hard to see why the victim's bad faith should work against her whilst the wrongdoer's bad faith should not hinder him, especially when the victim's harm is severe.

This point, it must be said, only applies to type-(2) rather than type-(4) cases. If the defendant agrees with that victim that she has been benefitted, any objection he has will not be in bad faith. But this concession does not change the overall result. Ultimately, the bad faith of the victim's claim is overridden by the paternalistic reasons to compensate her. Not only might she realise her error and be grateful for the opportunity to restore her welfare, but even if she does not she may spend the money on improving her wellbeing in some other way. This idea, which we can call *transferred compensation*, is a powerful reason to reject *The Objective View* in these cases. With this form of compensation, she can retain her conception of wellbeing but respond to the wrong in a way that both she and her injurer can recognise as genuinely promoting her wellbeing.

At this point it is worth returning to V's relatives. Another way that V's compensation can be transferred is by giving it to her family or dependants. Suppose that an action is brought posthumously by V's estate. It seems that this action should be granted notwithstanding that V does not believe she was harmed. V may well accept that the award of compensation will benefit *them*, and if they do not share her false view of wellbeing, they will recognise it as the right response to the harm V has suffered. This is similar to the case where V uses the money to improve some other aspect of her wellbeing. These alternatives show that it is too restrictive to think of legitimate compensation only as a response to a

particular harm inflicted by a wrong. Instead, compensation can be transferred either to improve some other aspect of V's wellbeing or to benefit those who bring an action on V's behalf.

6. Conclusion: Arriving at *The Mixed View*

I have made three general claims about the role of the victim's false beliefs in compensating for wrongful harm. First, when V is benefitted but believes she is harmed, we have autonomy-based reasons to compensate her. Failing to do this would not adequately respect the autonomy of a wronged victim when she is vulnerable, and cannot be justified to those who do not endorse the legally sanctioned theory of wellbeing. It may burden the defendant with a victim's mistake, but this can be justified, primarily because the defendant was in the best position to avoid being burdened in this way.

Second, although we should grant claims to victims in these cases, these awards must be limited to a reasonable threshold. It is not plausible that defendants should bear extreme burdens because of the erroneous beliefs of their victims. Nor does respect for autonomy require treating severely harmed victims, such as those who have suffered amputations or paralysis, equivalently to those with false, idiosyncratic beliefs. These implausible implications of *The Respect View* ought to be rejected.

Finally, when V is harmed but believes she is benefitted, our autonomy-based reasons to withhold compensation are overridden. The compensation can be transferred and spent on improving some other component of the victim's wellbeing, and there is nothing objectionably paternalistic in this since the award does not make it more difficult for her to pursue her chosen conception of the good. Again, this can be justified to defendants. They could have avoided liability by refraining from wrongdoing and they cannot complain about the bad faith of the victim if they also exhibit bad faith in relying on her mistaken conception to avoid paying damages.

In making these claims, I defended a compromise between *The Respect View* and *The Objective View* that I call *The Mixed View*. *The Mixed View* might be accused of arbitrariness in its attempt to strike this balance. Let me conclude by offering a few reasons why I think that any appearance of arbitrariness is purely superficial. One asymmetrical feature of *The Mixed View* is the way in which autonomy-based reasons operate. They are decisive in one type of case but overridden in another. This is because the strength of autonomy-based reasons partially trades on objections to paternalism. But the objectionableness of paternalistic acts seems to be largely conditional on violating a person's choices or making it more difficult for her to pursue her choices. On this plausible view about paternalism, the approach to compensation offered by *The Objective View* is only problematic in type-(1) rather than type-(2) cases. In the former, refusing to compensate V contradicts her choices. In the latter, offering compensation does not contradict her choices because the award is not forced upon her. Indeed, it is at her discretion that D's moral duty is enforced.

Secondly, in type-(1) and (2) cases, where V and D have different views about wellbeing, D's view is rejected and V is entitled to compensation. On the other hand, I also argued that V's compensation should not be calculated according to her sincere beliefs, but should be limited to a reasonable threshold. Is it not arbitrary to prioritise V's autonomy on the one hand but then limit compensation for D's sake on the other? On the contrary: this where the balance ought to be struck. It is reasonable to perceive a conflict of reasons in these cases. We have reasons to respect the autonomous choices of individuals, but we also have reasons to respond to real harm, and these reasons can conflict. *The Mixed View* shows greater sensitivity to the reasons we have, as *The Respect View* ignores our reasons to respond to harm and *The Objective View* ignores our reasons to respect autonomy.

On the basis of these considerations, I doubt that *The Mixed View* is arbitrary. Not only is it not arbitrary, but the discussion of this chapter also shows, I think, that *The Mixed View* is the most defensible and humane of the three approaches to compensation for wrongful harm canvassed here.

Corrective Duties

What principles tell us which corrective duties we ought to recognise and enforce? One popular answer is that corrective duties can be justifiably enforced when doing so minimises the costs of accidents and their prevention. Many, including myself, reject this vision of tort law, and must look elsewhere for a moral foundation for corrective duties, which does not rely for its justification on purely instrumentalist values.

Another type of theory is based on considerations of fairness. Since this categorisation is vague, many theories potentially fit this description. Here I want to focus on one whose central concept is *reciprocity* between victim and injurer. George Fletcher's theory of nonreciprocity provides a good starting point to investigate accounts of tort law which appeal to considerations of fairness rather than instrumental goals. Roughly speaking, it holds that a person's liability to pay compensation depends on whether she nonreciprocally imposed a risk of harm on others. In Fletcher's words, "a victim has a right to recover for an injury caused by a risk greater in degree and different in order from those created by the victim and imposed on the defendant – in short, for injuries resulting from nonreciprocal risks."¹

Fletcher's view is important for a number of reasons. First, it is founded on a principle of distributive justice, which guarantees equal security for individuals compatible with a like system for all,² and therefore is a rival non-instrumentalist account of tort law to corrective justice. This feature of the view will be applauded by those who deny the claim of some corrective justice theorists that tort law is insulated from issues of distributive or

¹ George Fletcher, "Fairness and Utility in Tort Theory", *Harvard Law Review*, 85(3), (1972), 527 – 573 at p. 542.

² See *ibid* at p. 550 and Jules Coleman, "Moral Theories of Torts: Their Scope and Limits: Part 1", *Law and Philosophy*, 1(3), (1982), 371 - 390 at p. 389.

political justice.³ Fletcher's article explicitly draws on the work of John Rawls, reinforcing the idea that the revival of non-instrumentalist thinking in political justice championed by Rawls and others also has implications for tort theory.⁴

Secondly, Fletcher's view has the potential to explain landmark decisions in the common law such as *Rylands v Fletcher*, in addition to the historical development of the relationship between negligence and strict liability. The wide interpretive potential of Fletcher's view is another reason to take it seriously as a moral foundation for liability. Even if there are doubts about the full scope of its interpretive ambitions, ascertaining its normative strength will help to determine whether and how the core practices and doctrines of tort law can be justified.

Third, Fletcher's theory has some initial plausibility. Reciprocity is an appealing concept, independently of its ability to explain the case law, and even those who are critical of Fletcher's work, such as Gregory Keating,⁵ have revitalised its central elements. In considering the nonreciprocity theory, we can also compare it to other candidate principles that purport to govern core cases of unintentional harm. Even if we eventually reject nonreciprocity, which I argue we ought to do, this comparative exercise will help us make positive steps towards a better theory of liability for unintentional harm.

³ See Ronen Avraham and Issa Kohler-Hausmann, "Accident Law for Egalitarians", *Legal Theory*, 12(3), (2006), 181 – 224; Larry Alexander, "Causation and Corrective Justice: Does Tort Law Make Sense?", *Law and Philosophy*, 6(1), (1987), 1 – 23, and Wojciech Sadursky, "Social Justice and Legal Justice", *Law and Philosophy*, 3(3), (1984), 329 – 354.

⁴ A number of theorists cite Fletcher's article as paving the way for a proper understanding of the normative basis of tort law. See for example Gregory Keating, "Reasonableness and Rationality in Negligence Theory", *Stanford Law Review*, 48, (1996), 311 – 385, at pp. 313 – 4 and Arthur Ripstein, *Equality, Responsibility and the Law*, Cambridge: Cambridge University Press, (1999).

⁵ See Gregory Keating, "A Social Contract Conception of the Tort Law of Accidents", in *Philosophy and the Law of Torts*, Ed. Gerald Postema, Cambridge: Cambridge University Press, (2001).

Let us now take a closer look at Fletcher's view. He claims that nonreciprocity explains doctrines of strict liability in tort. Take, for example, a claim under the rule from the Second Restatement of Torts that aeroplane owners and pilots are strictly liable for ground damage but not mid-air collisions.⁶ He explains this in virtue of the fact that the risk of ground damage is nonreciprocal: people on the ground do not impose risks on aircraft flying overhead but aircraft impose risks on them. Conversely, the risk of a mid-air collision is imposed reciprocally by all those who use air lanes. Nonreciprocity also *unifies* seemingly disparate pockets of strict liability. Liability for damage caused by falling aeroplanes or injuries inflicted by dangerous animals, as well as the rules laid down in *Rylands v Fletcher* and *Vincent v Lake Erie*, are motivated by the creation of a nonreciprocal risk.

Similarly, nonreciprocity accounts for the distinction between strict and fault-based liability. Fletcher notes that principles of negligence typically apply in contexts where people create and expose themselves to equivalent risks, such as driving on roads and playing in sports fields. In these contexts, the activity itself, as long as it is performed with due care, does not create nonreciprocal risks. If, however, the activity is performed negligently, the degree of risk increases, so negligent agents impose nonreciprocal risks relative to those who meet a reasonable standard of care. This explains why fault-based liability is the norm in these contexts: risks are only nonreciprocal when parties act negligently. Finally, Fletcher extends the analysis to intentional torts. An intentional blow or assault, he claims, "represents a rapid acceleration of risk, directed at a specific victim" which readily distinguishes "the intentional blow from the background of risk".⁷

Gregory Keating's work can be seen as a development of Fletcher's. One important criticism of Fletcher, which is made explicitly by Keating and implicitly by others, is that the distribution of harm matters more than the distribution of risk. This encourages Keating to reject Fletcher's view and adopt an alternative that advocates strict liability funded by

⁶ Fletcher, "Fairness and Utility in Tort Theory" at p. 548.

⁷ *Ibid* at p. 550.

compulsory liability insurance, designed to spread the costs of non-negligent accidents across all those who benefit from the activity.⁸ Keating's view has much to recommend it, but as will become clear, he conflates reciprocity with benefit, and these factors ought to be kept separate. Where relevant, I will refer to Keating's development of Fletcher, but I will mainly confine discussion to a critical analysis of the theory of nonreciprocity.

The primary aim of Fletcher's theory seems to be explanatory. It seeks to provide an account of the foundations of tort liability as it is, and stands or falls depending on how well it fits with actual cases and doctrines. It has, of course, been criticised from this perspective. For example, Jules Coleman doubts that nonreciprocity can explain intentional torts. A falling person 'rapidly accelerates' the risk to a person below, but unless she has deliberately fallen in order to apply force to another, she has not committed a battery (although Fletcher could respond that, on his scheme, the falling person would be excused because she was not in control of her actions).⁹ But criticisms of Fletcher's theory need not be based only on its explanatory capacity. As with many explanatory theories in law, the nonreciprocity principle does not merely seek to explain and unify tort practices, but to do so in a normatively favourable light.¹⁰ As I indicated in chapter 2, Coleman adopts a similar methodology in his own work, claiming that his approach provides "an explanation of our practices, or important parts of them, but explanations that make sense of the practice in the light of norms... that could withstand the test of rational reflection."¹¹

⁸ See Gregory Keating, "A Social Contract Conception of the Tort Law of Accidents".

⁹ Jules Coleman, "Justice and Reciprocity in Tort Theory", *Western Ontario Law Review*, 14, (1975), 105 – 118 at p. 115. Fletcher discusses excuses at pp. 551 – 6 of "Fairness and Utility in Tort Theory".

¹⁰ He claims that the principle is grounded in considerations of fairness and defends it by identifying normative problems with alternative utilitarian approaches. See "Fairness and Utility in Tort Theory" p. 550 and pp. 568 – 9.

¹¹ Coleman, *Risks and Wrongs* at p. 7.

I have expressed worries about this approach in more detail elsewhere.¹² In light of these, my interest in Fletcher's theory in this chapter is solely its normative power. That is, its relevance to the question of whether a person *morally ought* to be liable to compensate another. The moral basis of liability to compensate is, I take it, a fundamental issue of normative tort theory.

Fletcher assesses the paradigm of reciprocity against what he labels the paradigm of reasonableness, which represents "a rejection of non-instrumentalist values and a commitment to the community's welfare".¹³ Reasonableness is construed here as a balancing of costs and benefits in the manner of the famous Learned Hand test, and has obvious affinities with the economic approach to torts. Fletcher's concern with comparing his theory to instrumentalist alternatives means that he fails to address two non-instrumentalist considerations that merit attention. The first is the extent to which a risk can be *avoided*, either by the risk-imposer or the person on whom the risk is imposed, and the second is the issue of who *benefits* from the risk-imposing activity. The critique of the present chapter will ultimately show that Fletcher's theory ought to be rejected, but its purpose is not purely negative. In evaluating alternative principles grounded in avoidability and the distribution of benefits, I hope to make some progress towards a satisfactory account of the moral basis of liability to compensate.

Principles based on avoidability and benefit distribution have not been widely discussed in this context. It is more conventional to consider the fault principle, according to which liability depends on falling below an objective standard of reasonable conduct. I ignore fault in this chapter for a few reasons. First, strict liability seems justified in at least some cases. Anyone who shares this view will think that fault is not the whole story. Fletcher had the right idea by looking for deeper principles that apply to both fault and no-fault cases, and the benefit and avoidability principles are similarly broad in scope. Second,

¹² See chapter 2, section 3.

¹³ Fletcher, "Fairness and Utility in Tort Theory" at p. 542.

the force of the fault principle is derivative, at least in part, from the two principles I defend. The moral case for holding a negligent actor liable is substantially weakened if she acts solely or primarily for the benefit of a third party, and one reason why fault often justifies liability is that negligent harm is usually easier to avoid than non-negligent harm.

I begin by clarifying the three notions of nonreciprocity, avoidability and benefit. In particular, I analyse nonreciprocity into two distinct components: unequal risk distribution and asymmetrical causality of risk. In section 2 I briefly attempt to ascertain the moral underpinning of nonreciprocity, considering the concepts of fairness, standing and causation. Concluding that there is no compelling moral rationale for nonreciprocity, I proceed to the case analysis in sections 3 and 4. I argue that the two considerations of avoidability and benefit outweigh nonreciprocity and better explain our intuitive judgements about liability in core cases. In section 5 I evaluate the two components of nonreciprocity separately. I conclude that asymmetrical causality is insignificant. Unequal distribution of risk can be relevant, but is really a component of a larger principle or set of principles of distributive fairness, and these distributive concerns are not adequately captured by the principle of nonreciprocity.

1. The Three Principles

A. Nonreciprocity

Our first task is to clarify the three principles under investigation. The principle of nonreciprocity can be stated as follows:

Nonreciprocity Principle: Whenever A harms B, there is a stronger reason to let the cost fall on A than B if A imposed a nonreciprocal risk on B.

The idea of nonreciprocal risk-taking I employ here differs in certain respects from Fletcher's presentation, and these differences must be clarified. Fletcher claims that an injury resulting from a nonreciprocal risk is a necessary but not sufficient condition to impose a compensatory duty. A defendant may escape liability if her act is excused; if, for example, she was physically compelled to act or was unavoidably ignorant. But it does not follow that the loss ought to lie where it falls – merely that the defendant is not liable to pay. This leaves open the possibility that the victim has a claim for compensation against the community.

My definition of nonreciprocity brackets these issues. I do not directly consider excuses, in Fletcher's sense, as in all cases I examine the harm is foreseeable and agents are in control of their actions. As will become clear, I think that Fletcher's excusing conditions can be assimilated to the broader and more significant concept of avoidability. If a defendant is not in control of her body, this is indeed a good reason not to hold her liable, and this reason stems from the fact that she could not have avoided causing the injury. Moreover, I am mainly concerned with two party cases where loss spreading is not an option. The possibility of loss spreading introduces new and complex considerations, and we ought to begin by determining which principles are most defensible in two party cases, before asking how these principles are modified when loss spreading is an option. Effectively, then, there are only three possible outcomes in the cases I will consider: the injurer compensates the victim, the victim bears the loss, or the costs are divided between them.

Second, we must distinguish the degree of risk from the magnitude of potential harm. In his discussion, Coleman, although he ultimately rejects nonreciprocity as a general theory of tort liability, incorporates both concepts by defining nonreciprocity in terms of *expected* value or disvalue. The expected value of a risk takes into account both the magnitude of a potential harm and the likelihood of its occurrence.¹⁴ To avoid complications involving the distinction between these two relevant factors, I will adopt this formulation.

¹⁴ Coleman, *Risks and Wrongs* at p. 254.

Third, as Coleman has pointed out,¹⁵ Fletcher's criterion is ambiguous between the size and the social tolerability of risks. In this paper I will exclude the latter, as I agree with Coleman that the social tolerability of a risk has little to do with nonreciprocity, but is rather an indicator of social value.¹⁶ If a risk-creating activity is socially tolerated, this suggests that it is valued by the community. The fact that an activity is socially valued implies that a large group of people benefit from it in some way, and may be a reason against liability if those who undertake it cause harm. But if so, this reason does not stem from nonreciprocity. The activity may involve the same degree of nonreciprocal risk as some other activity that is hardly valued at all. The two types of activity are instead distinguished by the fact that the former benefits a larger group of people. I will outline an alternative principle that rests on the distribution of benefits produced by an activity in section 1C.

Fourth, Fletcher claims that nonreciprocity should be measured against a 'community of risk takers' rather than a single party. Thus, "keeping a domestic pet is a reciprocal risk relative to the community as a whole; driving is a reciprocal risk relative to the community of those who drive normally; and driving negligently may be reciprocal relative to the even narrower community of those who drive negligently."¹⁷ The notion of a 'community' of risk-takers also needs refining. On Fletcher's view, careful drivers constitute a community of reciprocal risk-takers in virtue of their competent participation in a shared activity. But this mutual participation and adherence to a reasonable standard of care does not guarantee that risks will be reciprocal. To see this, consider the following:

Minis and Hummers: Through no fault or choice of their own, Alice and Rachel are only able to retrieve clean water by car once a week. There are only two modes of transport: Minis and Hummers. Due to their weight and strength, Hummers impose

¹⁵ Coleman, "Justice and Reciprocity in Tort Theory" at p. 108.

¹⁶ Ibid at p. 111.

¹⁷ Fletcher, "Fairness and Utility in Tort Theory" at p. 549.

much higher risks on pedestrians and have the potential to cause more harm than Minis. Both Alice and Rachel drive to obtain water each week, but Alice chooses a Mini while Rachel chooses a Hummer.

Both Alice and Rachel fall within Fletcher's definition of a community of reciprocal risk-takers as they meet a reasonable standard of care on the road. Nevertheless, Hummer-drivers such as Rachel impose nonreciprocal risks on Mini-drivers such as Alice because of their decision to drive a more dangerous vehicle. Fletcher could respond that reciprocity only requires rough equivalence of risk, rather than precise equivalence. It is true that, if reciprocity is morally important, it plausibly allows for negligible differences. But in any community of risk-takers, it remains a possibility that risk distribution is sufficiently different to undermine the assumption that the individuals in that group impose equivalent risks on one another. At best, this sense of 'community' is a useful generalisation.

There are, however, two other important meanings of 'community' implicit in Fletcher's formulation, which must also apply to ours. First is the sense in which the risks that one person imposes on others may still be reciprocal even though they are not reciprocal *on a given occasion*. Imagine that in *Minis and Hummers*, both Alice and Rachel choose Hummers and drive with equal frequency. On one occasion, Alice chooses to walk and Rachel imposes a risk on her as she cruises past in her Hummer. Even though the risks are nonreciprocal in this instance, Alice falls within the relevant community as she imposes equivalent risks at other times.¹⁸ Second, it does not necessarily matter on *whom* the risks are

¹⁸ There is a complication here. Two risk imposers, A and B, may not be in a morally equivalent position merely because they impose equal average risks over a designated period of time. Suppose Alice walks 99 days out of a hundred and on the 100th day drives a Hummer at 150 mph. Rachel drives a Mini carefully every day. Even if they impose equal average risks over the course of 100 days, Alice's behaviour is still impermissible while Rachel's is permissible. This suggests that we

imposed. The relevant form of reciprocity is relative to the average risk imposed by someone who participates in a given activity, regardless of who is subject to the risk.

Having made these initial clarifications, let us define nonreciprocity more precisely. We can see that nonreciprocity usually has two components. The first pertains to the *distribution* of the risk. Specifically, nonreciprocity requires an unequal distribution, whereas reciprocity implies a more or less equal distribution. The distribution of risk between A and B will be unequal if A is subject to a risk but B is not, or if A and B are both at risk, but A is subject to a much lower risk than B (as in *Minis and Hummers*). The second component is *asymmetrical causality*. Nonreciprocity between A and B will also be established if A causes a risk to B and B does not cause a risk to A. Note that the two conditions are not jointly necessary for nonreciprocity. As *Minis and Hummers* shows, if A imposes a risk, X, on B, and B imposes a lower risk, X-10, on A, it is still a case of nonreciprocity despite the absence of asymmetrical causality. Generally, we should be aware that the two components are distinct. We can imagine cases in which risks are distributed *unequally* between two people and *neither* of them are a causal source of the risk (unequal risk but not asymmetrical causality). Similarly, we can imagine cases in which risks are distributed *equally* between two people and *one* of them is the causal source of the risk (asymmetrical causality but not unequal risk). Despite their distinctness, both components capture important aspects of nonreciprocal risk-taking.

B. Avoidability

Now that we have clarified the principle of nonreciprocity, we must similarly clarify the two other considerations that are plausible determinants of liability, and that may compete with

cannot straightforwardly aggregate risks. In what follows I will ignore issues of aggregation as they do not impact upon the basic moral relevance of reciprocity.

or supplement nonreciprocity. The first of these is avoidability. We can set out an avoidability principle as follows:

Avoidability Principle: Whenever A harms B, there is a stronger reason to let the cost fall on the party who had a better opportunity to avoid the risk.

In order to explicate this principle, we must answer some preliminary questions: what different types of avoidability can be distinguished? Which are normatively relevant to liability to compensate? What does it mean for one person to have a ‘better opportunity’ to avoid a risk?

We can begin by distinguishing fact-relative avoidability from belief and evidence-relative avoidability.¹⁹ Some harm, H, is fact-relative avoidable for an agent, S, if and only if S could avoid it if she knew all the relevant facts. H may be fact-relative avoidable even if S does not know and could not know that she could avoid it, as it would still be true that S could avoid H *if* she knew all the facts. The physician who gives a treatment to a patient and kills her due to an unforeseeable and undiscovered allergic reaction could have avoided the death only in the fact-relative sense. Conversely, harm is belief-relative avoidable for S if and only if S can avoid it given S’s beliefs, and evidence-relative avoidable if and only if it is avoidable given the evidence available to S, about which S ought to be aware. The death would be belief-relative avoidable for the physician if she knew that a severe allergic reaction is possible in some patients, and evidence-relative avoidable if she ought to possess this knowledge.

¹⁹ See Derek Parfit, *On What Matters*, Vol. 1, Oxford: Oxford University Press, (2011), pp. 150 – 62.

See also Victor Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law*, Oxford: Oxford University Press, (2011), pp. 219 – 20.

Mere fact-relative avoidability gives us no reason to impose liability. The harm must be foreseeable given the agent's knowledge or the knowledge she ought to have.²⁰ Otherwise it makes no sense to say that *the agent*, who is not an omniscient being, could have avoided the outcome. This, of course, is the position in English tort law. *Roe v Minister of Health* affirmed that the duty of care in negligence owed to medical patients does not include taking precautions against risks that are not foreseeable given the prevailing scientific knowledge of the time.²¹ The foreseeability requirement also constrains the extent to which the avoidability principle should take into account the agent's past opportunities to avoid the outcome. A past opportunity to avoid an outcome only counts in the relevant sense if it is roughly within the realm of belief or evidence-relative avoidability. It must at least be a foreseeable consequence of some avoidable act that it may place the agent in a position where she unavoidably performs some *further* act that risks harm to others.²²

There is another distinction between two broad categories of avoidability. The first can be called *factual avoidability*. If Rachel is pushed into a victim, effectively being used as a projectile, she is unable to avoid the threat that she poses in this sense. If Alice suffers from a condition that makes her lose control of her limbs, the punch she throws at a bystander is also unavoidable in the factual sense. Again, this form of avoidability is subject to the agent's prior opportunities. If Rachel knows that she will be pushed into the victim but has conspired for this to happen, the harm is still avoidable even though she cannot avoid it

²⁰ I am glossing over the difficulty of holding an agent responsible for her ignorance. For an elaboration of these difficulties, see Michael Zimmerman, "Moral Responsibility and Ignorance", *Ethics*, 107 (3), (1997), 410 – 426.

²¹ [1954] 2 All ER 131.

²² Similar points have been made in the literature on luck egalitarianism. See Peter Vallentyne, "Brute Luck, Option Luck, and Equality of Initial Opportunities", *Ethics*, 112(3), (2002), 529 – 557, at pp. 532 – 3.

once she is pushed.²³ Finally, in these two cases the harm is *more* avoidable for Rachel than for Alice. This implies that avoidability is a matter of degree, determined by the costs involved in avoiding the outcome. It is more costly for Alice to avoid the harm because she must forgo a permissible activity that is instrumentally useful, namely appearing in public. The costs to Rachel in avoiding the attack consist in not conspiring to harm the victim. Morality already requires this of her, so it cannot be considered a relevant cost.²⁴

The second type can be labelled *normative avoidability*. Even if harm is factually avoidable, it may be unavoidable without some normative cost. In other words, an agent might have good *reasons* to impose a risk or cause harm. An agent may even have a *duty*, all things considered, to do something that will cause harm. Suppose Rachel is Alice's bodyguard and has a duty to protect her. When Alice is attacked by Victoria, Rachel has a duty to harm Victoria to defend Alice, assuming the defensive harm is proportionate and necessary to avert the threat. The harm Rachel has a duty to inflict is factually avoidable. Nevertheless, it is unavoidable in the sense that Rachel could not have avoided it without breaching her duty to Alice. Similar claims apply to a variety of other cases. For example, if ambulance drivers have a moral duty to respond promptly to emergency calls, the risks they impose on road users in arriving at the scene are unavoidable in this normative sense.

Normative costs may be moral, as in the case of Alice and the ambulance driver, or prudential. In *Hummers and Minis*, Alice and Rachel would suffer a huge personal cost if

²³ Again, this is reflected in the law of torts. A person who causes negligent harm to others as a result of a condition that causes her to temporarily lose control of her body will only be liable if she has prior warning that she suffers from such a condition. See *Mansfield v Wheetabix Ltd.* [1997] EWCA Civ 1352 and *Roberts v Ramsbottom* [1980] 1 WLR 823.

²⁴ On this point, see Thomas Scanlon, "Thomson on Self-Defense" in *Fact and Value: Essays on Ethics and Metaphysics for Judith Jarvis Thomson*, Eds. Alex Byrne, Robert Stalnaker and Ralph Wedgwood, Cambridge, MASS: MIT Press, (2001). See also Victor Tadros. *The Ends of Harm: The Moral Foundations of Criminal Law*, Oxford: Oxford University Press, (2011) at p. 232.

they did not impose risks on others by retrieving water. Agents clearly have strong prudential reasons to protect their wellbeing, and if an agent cannot avoid some harm or risk without incurring a significant personal cost, it is normatively unavoidable in a prudential sense.

Normative costs vary in significance. The cost of breaching a duty is greater than that of forgoing an opportunity to pursue a morally good end. Suppose Alice wants to volunteer at the local homeless shelter and in doing so must drive across town, thus imposing modest risks on road users. She does not have a duty to volunteer, but it is a morally good end that does not merely further her own interests. The risks she imposes are unavoidable without forgoing an opportunity to pursue this end, but the risks imposed by the ambulance driver are more unavoidable given that she has a duty to act, rather than just a positive moral reason. Both factual and normative avoidability engage the *Avoidability Principle*, but for simplicity I will only consider cases where avoiding some act involves a personal cost in the remainder of this chapter.

If avoidability is influenced both by the cost of avoiding acts and the reasons one has to undertake them, avoidability can be a matter of degree. It can vary according to the magnitude of the costs or the strength of the reasons. One might think that we should simply speak about reasons. After all, if some agent will incur a cost unless she imposes a risk, surely she has a reason to impose the risk which is proportionate to the size of the cost. But even if this is true of prudential reasons, it is not true of moral reasons, so we should not conflate costs with reasons. The costs to two agents in not Φ -ing may vary significantly even when both have decisive reasons not to Φ . This is true of the saint and sinner who each have decisive reasons not to murder. The saint incurs minimal cost – perhaps even a benefit – in acting on her decisive reasons, as she is already predisposed to fulfil her duties gladly. The sinner has a strong craving for violence, and feels herself frustrated and unfulfilled if she does not succumb to her urges. Clearly, the sinner follows her decisive reasons at greater cost to herself, but it does not follow that she has less reason to refrain from killing.

Cravings make a difference to personal costs and the prudential reasons that stem from them, but they do not have the same effect on moral reasons or duties.

A final complication in the concept of normative costs should be mentioned. When evaluating avoidability we should take into account disjunctive reasons to perform one of a number of acts that impose uniform risks, rather than just reasons to perform single risk-imposing acts. If Alice must choose between two acts, A and B, which impose equal risks, on pain of suffering cost X, she has a disjunctive reason to choose one or the other which is as strong as the reason she would have to perform a single act on pain of suffering X. This is true despite the fact that, in the first example, both A and B are entirely avoidable when considered individually. This may make any avoidability principle difficult to operate in practice. Most drivers can avoid operating their dangerous vehicles on any given occasion, but over the course of a lifetime it is almost inevitable that one will need to drive. The same point can be stated in terms of reasons. A driver does not have a strong reason to get into her car on each separate occasion, but she has very strong reasons to drive at some point, as otherwise she will be unable to secure basic goods.

C. Benefit

Risk-imposing activities are often beneficial and this may affect liability to compensate for harm caused when those risks manifest. The benefits of risk-imposing conduct will often accrue to the person who imposes the risk, but this is not always the case. Since the benefit of an activity is a distinct factor with potential moral relevance, as with the two previous considerations, we must make a few preliminary remarks to clarify the arguments to follow. We can begin by explicitly stating a benefit principle.

Benefit Principle: Whenever A harms B, there is a stronger reason to let the cost fall on the party that benefits more from the risk-creating activity.

Benefits may be of different types and amounts. Some types of benefit may have priority over others, so that they affect liability even when the beneficiaries are the same. If a public body imposes a risk on citizens to provide them with clean water, then that is a stronger reason against holding it liable in the event of an injury than if it imposed the same risk for the sake of a public garden. This is because the type of benefit secured by clean water has priority over the aesthetic benefit of a public garden. Similarly, all else being equal, the greater the *magnitude* of the benefit, the easier it is to justify imposing the risk. If there are two possible sanitation systems a public authority could implement, it might be justified in imposing greater risks for the sake of one over the other if it is significantly more safe and efficient in the long term.

It is also crucial to identify the beneficiaries of the activity. There are a number of possibilities here. An activity may benefit the person who imposes the risk, the person on whom the risk is imposed, a particular third party, or a group of people, either including or not including those who are party to the risk. For example, if a person imposes a nonreciprocal risk on another by operating heavy machinery in the course of a construction project, they may do so for their own benefit and for the benefit for a third party (their employer). If a person imposes a risk in the course of building a public hospital, a group of people (most of those who will be treated there) are potential beneficiaries of the activity. Finally, if a surgeon operates on a patient, the person on whom the risk is imposed is the prime beneficiary. As these brief remarks show, there are many benefits that risky activities generate and many ways they can be distributed, meaning that the *Benefit Principle*, like the *Avoidability Principle*, will sometimes be difficult to apply in practice. Nevertheless, there are some cases, such as medical procedures, where it is clear who benefits most significantly from the risky activity.

2. The Moral Underpinning of Nonreciprocity

The *Nonreciprocity Principle* has some immediate appeal. After all, if A imposes a nonreciprocal risk on B, B is the one who stands to incur the greatest loss. A is either not subject to a risk at all, or is subject to a much lesser risk. Moreover, even if A and B are subject to the same risk, perhaps it is troubling if A causes the risk to B but B does not cause the risk to A. The two components of nonreciprocity – unequal risk distribution and asymmetrical causality – have some ostensible moral relevance.

Fletcher believes that his view is based on a principle of fairness analogous to Rawls' first principle of justice, which is that everyone has a right to a maximum degree of security compatible with equivalent security for all.²⁵ But there is reason to doubt whether nonreciprocity is really a principle of fairness. Consider the distributive component of the principle: inequality of risk. An unequal distribution of risk may afford the victim a right to recover compensation. But unequal risk distributions are not necessarily *unfair* (this would assume that strict equality is the correct criterion of distributive fairness). I elaborate on this problem and related difficulties in section 5, so I will say no more about it here. In any case, distributive fairness can only account for one element of nonreciprocity. Even if unequal risk distribution is unfair, the moral significance of asymmetrical causality remains mysterious.

Another possibility is that nonreciprocity is intuitively supported by the significance of causation. The significance of causation is controversial, and there is no space to investigate it in any detail here, but two doubts can briefly be stated. First, even if pure causation is morally relevant in cases where a victim is harmed, it is difficult to see how it could be relevant when applied exclusively to risk. One of Coleman's criticisms can be interpreted along these lines. He argues that nonreciprocity is implausible because it implies that if A and B negligently impose risks on one another and A is injured, A cannot claim

²⁵ Fletcher, "Fairness and Utility in Tort Theory" at p. 550. For doubts about the relationship between Fletcher's principle and Rawls', see Coleman, "Justice and Reciprocity in Tort Theory" at pp. 117–8.

against B. Yet in this event the law will permit A to recover from B.²⁶ If this is the correct view, morally as well as legally, then the fact that B causes *harm* is decisive. This outweighs the fact that *both* A and B cause risks. Another way to state the point is that the *Nonreciprocity Principle* implies that A's entitlement to claim from B can depend on whether B causes an *unrealised* risk to A. But why should this matter?

Moreover, the suggestion that causation provides the moral attraction of nonreciprocity is vulnerable to the same objection as inequality of risk distribution: it only accounts for one component of the principle. If part of the intuitive appeal of the principle is that it offers remedies for the consequences of unequal risk distribution, this can be maintained in the absence of causation. Risks can be distributed fairly or unfairly amongst a group of people even if none of them causes the risk. A third party could unfairly distribute a risk between A and B, as a result of which A is injured. If the third party cannot be traced or is insolvent, and we can only distribute *ex post* costs between A and B, the unfairness of the initial risk distribution may give us reasons to hold B liable to some degree (this would not be a tort claim, of course, but considerations of fairness such as this are relevant to the more fundamental question of the moral basis of liability to compensate. An action in tort law is not the only way compensation can be transferred from one person to another). If so, this claim for compensation cannot be grounded in asymmetrical causality, since neither A nor B caused any risk. This shows that even if causation is relevant here, it cannot be the moral rationale for *both* elements of the principle. The distinctness of the two components makes it difficult to see how there can be a coherent moral underpinning for the whole principle.

There is another possibility that promises to provide this rationale. The general idea is that even if A is responsible or at fault for harming B, A cannot claim from B if A lacks *standing* to do so. There are a number of reasons why one party might lack standing to seek recompense from another, and the claim we are considering is that reciprocal risk-imposition is one such reason. One advantage of this claim is that it appeals to the absence of *both*

²⁶ Coleman, *Risks and Wrongs* at p. 268 – 9.

components of nonreciprocity. That is, it depends on both parties imposing a risk *and* that risk being equal. Both asymmetrical causality and unequal risk distribution are *prima facie* the sort of asymmetries that might undermine the standing of one party to claim from another. Perhaps, then, the notion of standing will provide a unified rationale for the *Nonreciprocity Principle*.

To assess the strength of this argument, we can begin by following a distinction between two ways in which one person might lack standing to complain about the actions of another.²⁷ First, A may be partially responsible for B's wrong – the very wrong about which A complains. The extent to which A is at fault might vary between creating the conditions in which B can commit the wrong to overtly encouraging her to do so. B's act is still wrong, and she may be answerable to others, but not to A, who lacks the standing to call her to account. This argument is inapplicable in the present case since standard instances of nonreciprocity do not require that A has any responsibility for B's risky activity at all. The other way in which a person may lack standing relies on the notion of hypocrisy. If A has committed the *same wrong* as B, it would be hypocritical of her to complain about B's wrong. In cases of reciprocal risk-taking one party will have imposed the same risks as the other, so perhaps lack of standing through hypocrisy explains the appeal of the *Nonreciprocity Principle*.

Unfortunately, this is a weak rationale. It only gives a negative reason against imposing liability on someone when a reciprocal risk is imposed on her, rather than a positive reason to impose liability on someone who imposes a nonreciprocal risk on others. It is hardly a positive reason in favour of liability that one party would not be a hypocrite to

²⁷ See Victor Tadros, "Poverty and Criminal Responsibility", *Journal of Value Inquiry*, 43, (2009), 391 – 413; Anthony Duff, "Blame, Moral Standing and the Legitimacy of the Criminal Trial", *Ratio*, 23, (2010), 123 – 40; and G. A. Cohen, "Casting the First Stone: Who Can, and Who Can't, Condemn the Terrorists", *Royal Institute of Philosophy Supplements*, 81, (2006), 113 – 136.

demand compensation from the other. The *Nonreciprocity Principle* should offer us a positive reason for liability.

But there are bigger problems than this. Returning to Coleman's example, we can see that even the negative reason is not always decisive. When two negligent motorists impose reciprocal risks on one another but one of them is injured, the injured party still has a claim against the other. It is no defence for the injurer to say that the victim has also negligently imposed risks. The actual outcome – that one party has been injured by another's wrongdoing – is the greater concern, and overrides the hypocrisy of the victim's claim.

Finally, those who argue that hypocrisy can sometimes undermine standing tend to focus on criminal or serious moral wrongdoing. Hypocrisy is more worrying in these cases than in tort law, where liability either flows from less significant wrongs, such as negligence, or even the consequences of permissible acts. Suppose Alice and Rachel own neighbouring land. Alice stores large amounts of water while Rachel keeps dangerous animals. Each takes all necessary precautions to ensure that these dangerous objects do not escape and cause damage, but each still imposes some risks on the other. If Alice's water escapes and drowns Rachel's animals, is it really true that Rachel's claim for compensation is hypocritical? Alice and Rachel may have imposed equal risks, but they did so in the course of a permissible activity.

This fact seems to undermine the hypocrisy argument. Victor Tadros explains the force of the argument in the following way: hypocrisy rests on the idea that the reasons against an action, or the reasons that make it wrongful, apply to one party but not the other. The hypocrite implies that the moral standards against which she measures another do not apply to her. This is erroneous because the very idea of a reason for action is neutral between agents.²⁸ But this asymmetry depends on a wrongful act, and cannot apply in the present case, because Rachel acts permissibly; there were no decisive reasons for her not to keep dangerous animals. Rather, her claim depends on her having suffered an adverse

²⁸ Tadros, "Poverty and Criminal Responsibility", p. 396.

outcome. She can claim this and consistently hold that she would be similarly liable if her animals had escaped and Alice suffered an adverse outcome. Her claim for compensation simply does not imply that she holds herself to different moral standards.

Perhaps there is another compelling rationale for nonreciprocity. I find it difficult to see what that would be. Whether A imposes an unrealised risk on B simply seems irrelevant in itself to whether A is entitled to claim from B.

3. Nonreciprocity and Avoidability

Having failed to find a compelling rationale for the *Nonreciprocity Principle*, we can now investigate whether it has plausible implications when compared to other principles. How significant is nonreciprocity relative to the other two principles – the *Avoidability Principle* and the *Benefit Principle*? Fletcher argues that it justifies the victim's right to obtain compensation in a wide variety of tort cases. This claim is false if the principle lacks normative power when compared to the other principles that we are considering here. To see whether this is so, we can begin by comparing nonreciprocity with avoidability. The *Avoidability Principle* states that if A harms B, we have a stronger reason to let the cost fall on the party who had a better opportunity to avoid the risk. What we will see is that it is significant whether one party could have avoided either imposing the risk on another person or having the risk imposed on them. This is true regardless of whether there is reciprocity or nonreciprocity of risk imposition. This suggests that avoidability is a more important factor for liability than nonreciprocity.

Consider:

Trolley Island: Alice and Rachel live on a small island in which there is a user-operated trolley that transports inhabitants to a well, passing through a popular foraging site, imposing foreseeable risks on others. Alice is a trolley-operator and

imposes a nonreciprocal risk on Rachel, a forager. Alice must operate the trolley to retrieve water otherwise she will suffer a significant personal cost, as she has no other source of clean water. She will experience considerable pain and reduced mobility as a result of her failure to get clean drinking water. Rachel will not suffer a similar cost if she does not forage there as she has another source of food – she is merely foraging for decorative flowers. Alice puts up signs, warning Rachel of the risk, to ensure that she has an adequate opportunity to avoid it. Alice operates the trolley and Rachel is injured.²⁹

In *Trolley Island*, the *Avoidability* and *Nonreciprocity* principles yield conflicting results. The former implies that the loss should lie where it falls, while the latter implies that Rachel has a valid claim for compensation against Alice. Intuitively, the *Avoidability Principle* yields the correct result. The nonreciprocity of the risk seems largely irrelevant here, given the difference in avoidability. It would be harsh to expect Alice to avoid imposing the risk given that the alternative carries a serious cost, and it is reasonable to expect Rachel to bear the cost given that she could easily have avoided being subject to the risk.

It might be objected that avoidability only overrides nonreciprocity in *Trolley Island* because Alice and Rachel do not live in a state that bears obligations to distribute resources justly. Despite deep disagreement about what justice requires many views will hold that Alice would not suffer such a significant cost if she had a fair share of resources. Distributive justice, at the very least, requires access to basic necessities such as clean drinking water.

However, this is not a decisive objection. We can demonstrate this by modifying the case so that Alice has a better opportunity to avoid imposing the risk. We can imagine that Alice will only suffer a moderate cost – she will need to obtain water elsewhere, which

²⁹ Arguments from avoidability along these lines were first made by Thomas Scanlon. See *What We Owe to Each Other*, Cambridge, MA: Harvard University Press, 1998), pp. 256 – 267.

requires a non-negligible expenditure of time and money that could be used to further her projects or take care of her family. In avoiding the risk, she does not suffer so severely that she falls below a threshold that she is entitled to be kept above as a matter of justice. Nevertheless, the cost she suffers is not insubstantial, and as before, Rachel is merely foraging on a whim for decorative flowers. Even here, the *Avoidability Principle* yields the intuitive result. Rachel has a much better opportunity to avoid being subject to the risk than Alice has to avoid imposing it, and so the loss should be borne by Rachel. At the very least, this case shows that Rachel is not entitled to *full* compensation as suggested by the *Nonreciprocity Principle*.

A quick digression into legal defences is worthwhile here. Note that I am not arguing that Rachel has voluntarily assumed the risk. She may not assume the risk even if she has knowledge of it: a voluntary assumption requires acquiescence and this does not follow from mere knowledge.³⁰ To be sure, if the opportunity to avoid a risk is sufficiently large and one fails to take it, voluntary assumption may be *inferred* from one's behaviour, but the relevance of avoidability goes beyond these cases. One can have a sufficient opportunity to avoid a risk even if one fails to follow the warnings or even notice them. In this event one cannot be said to have acquiesced to the risk, but one's opportunity to avoid it is still a reason against imposing liability on the injurer.

Trolley Island is also similar to cases of contributory negligence. One important difference is that it takes into account the *injurer's* opportunity to avoid imposing the risk as well as the victim's opportunity to avoid bearing it. Moreover, in appealing to the concept of avoidability, it is also less misleading than contributory negligence. One reason for this is that contributory negligence is not really negligence at all, at least as that concept is otherwise understood. Negligence implies wrongdoing, but in most cases of contributory negligence we would not say that the claimant has acted wrongly. A finding of contributory negligence may simply reflect the agent's opportunity to protect herself from potential harm,

³⁰ Again, the position in English tort law is similar. See *Smith v Baker & Sons* [1891] AC 325.

which she failed to take, even if in doing so she acted permissibly. The fact that one is permitted to take greater gambles with one's own physical safety than with the safety of others suggests that contributory negligence rarely involves wrongdoing. Its title in law is therefore something of a misnomer (this is further demonstrated by the requirement that negligence, unlike *contributory* negligence, presupposes the existence of a duty of care). These comments suggest that the concept of avoidability underlies many of our intuitions about defences in tort. But this is a subject for discussion in its own right, so I will dedicate no more time to it here.

4. Nonreciprocity and Benefit

In *Trolley Island* we found that the *Avoidability Principle* overrides the *Nonreciprocity Principle*. But suppose we vary the example so that Alice and Rachel have an equal opportunity to avoid the risk (perhaps they can both avoid it at minimal cost to themselves), rendering the *Avoidability Principle* inapplicable. Alice operates the trolley, imposing a nonreciprocal risk on Rachel, as a result of which Rachel is injured. In this variation, since both parties can equally avoid the risk, perhaps we have stronger reasons to impose liability on Alice than we have to let the loss lie where it falls, based on the fact that the risk is nonreciprocal. This conclusion is too quick, however, as the judgement that Alice is liable might issue from the alternative ground that she is the beneficiary.

In Fletcher's original article, he only considers cases in which the risk-creating activity benefits the risk-imposer more than the person on whom the risk is imposed, such as injuries caused by crashing aeroplanes, damage done by wild animals, blasting, and so on. Others, including some critics of Fletcher's theory such as Gregory Keating, conflate nonreciprocity of risk with the thought that the risk-imposer is the prime beneficiary. Keating seems to base this assumption on his prior claim that when risks are reciprocal each party relinquishes an equal amount of freedom and gains an equal amount of security, thus

creating mutual benefit.³¹ But irrespective of whether this prior claim is true (and there is reason to doubt that it is; it will depend on how best to conceptualise freedom and security), the conflation of nonreciprocity with benefit is still a mistake. There are a plethora of benefits to take into account, other than the forms of security and freedom that are produced by being free from risks and possessing the right to impose risks respectively, and these may be distributed in a variety of ways. As I clarified in section 1C, there are many cases where one party is the nonreciprocal risk-imposer, but other parties benefit to an equal or greater degree. We should therefore formulate the notion of benefit into an independent principle to assess cases where it conflicts with the *Nonreciprocity Principle*.

The *Benefit Principle* outlined in section 1C implies that if A harms B, there is a stronger reason to let the cost fall on B than A if B benefits more from the risk-creating activity. If this is true, despite A imposing a nonreciprocal risk on B, the *Benefit Principle* overrides the *Nonreciprocity Principle*. Admittedly, there are some cases where the *Benefit Principle* does not give correct results. If Alice imposes a high risk of injury on Rachel to win Rachel some jelly beans, the fact that Rachel benefits more than Alice will not cancel her claim for compensation if she is injured. The mere fact that Alice enjoys a greater benefit than Rachel, however slight, cannot justify the imposition of any risk, however great. The *Benefit Principle* only comes into its own when the benefit to the victim is sufficiently large. But there are a number of common cases in which this is true. When a surgeon operates on a patient with reasonable skill, she imposes a nonreciprocal risk on him. If the patient is injured by the procedure, he has no moral claim against the surgeon because the risk-imposition is primarily for his benefit, and this benefit is sufficiently large to justify the risk.

This case involves a complication because patients usually consent to treatment.³² Consent may outweigh the relevance of benefit, because usually a person has no claim

³¹ Gregory Keating, "A Social Contract Conception of the Tort Law of Accidents", at p. 31.

³² In a footnote Fletcher agrees that injuries in the course of consensual, bargaining relationships pose special problems. However, he does not explore or recognise the extent of these difficulties, or

against someone who imposes a *consensual* risk on her, even if she does not benefit from it. If Alice asks Rachel to slap her in the face in an act of deliberate self-punishment, Alice has no claim against Rachel for any subsequent harm. Nevertheless, acting for the benefit of another³³ by imposing a nonreciprocal risk on them can still cancel a moral claim for compensation in the absence of consent if consent cannot reasonably be sought.

Consider:

Unconscious Patient: A doctor happens upon an unconscious patient, who is in danger of death unless the doctor acts quickly. There is no one around who might give consent to any treatment on behalf of the unconscious person. The doctor can perform a procedure that will probably save the patient from death, but also carries a small risk of paralysis. She performs the procedure, saving the patient's life, but unfortunately the patient also loses the use of his right arm.

In *Unconscious Patient*, although the doctor imposes a nonreciprocal risk on the patient to which he does not consent, he has no claim for compensation against her because she acts for his benefit.

Claims based on nonreciprocity are also undermined when the activity has a broader social benefit that affects the victim. The victim need not be the sole beneficiary, nor need she be benefitted directly. Even if a pedestrian who does not drive is subject to nonreciprocal

separate consent from benefit. It is therefore unclear whether he accepts that benefit is morally relevant independently of consent. See "Fairness and Utility in Tort Theory" at p. 548 n. 43.

³³ In acting for the benefit of another, it is sufficient that one acts for their *prospective* benefit. A surgeon is not liable to compensate a patient if she acts for the patient's benefit, even if her efforts are not successful. On this point, see also Victor Tadros, *The Ends of Harm* Ch. 10 and Jeff McMahan, "The Just Distribution of Harm Between Combatants and Noncombatants" *Philosophy and Public Affairs*, 38(4), (2010), 342 – 379 at p. 360.

risks on the road, it is still implausible that she has a claim against a driver who injures her without fault (although she may have a claim against the wider community). Non-drivers are typically benefitted in numerous ways by a general permission to drive (emergency responses, delivery of goods and services, and so on), and this undermines their claims against risks inherent in driving.³⁴

Arguably, it is not even necessary that the pedestrian is personally benefitted, as long as a sufficient benefit accrues to others to outweigh the pedestrian's complaint about being subject to nonreciprocal risk. An effectively identical case is discussed by Aaron James in his analysis of contractualism and tort theory.³⁵ In this case, an Amish farmer rejects the general benefits of social life but lives underneath a flight path. The aircraft flying overhead impose nonreciprocal risks on the farmer. This seems unfair, as the farmer does not enjoy the social benefits of aviation, but on the other hand, many more deaths would result if air travel were impermissible. James applies this argument to *ex ante* contractualism, but similar claims can be made about liability to compensate. If the Amish farmer is injured she may have a claim against the wider community, but the widespread benefits of aviation, even if they do not accrue to the victim, count against a claim for compensation from the aviation company. Nevertheless, we ought to be careful about how we characterise the source of the benefit. Strictly speaking, the benefits under discussion (transportation of medicine, for example) only require commercial air travel. Would the situation be different if the farmer was injured by a privately owned aircraft? In principle, this case, in which the activity is undertaken for private benefit, should indeed to be distinguished from those in which the activity benefits society at large.

³⁴ See also Gregory Keating, "A Social Contract Conception of the Tort Law of Accidents" in *Philosophy and the Law of Torts*, Ed., Gerald Postema, Cambridge: Cambridge University Press, (2001) at p. 29.

³⁵ Aaron James, "Contractualism's (not so) Slippery Slope", *Legal Theory*, 18(3), (2012), 263 – 292 at pp. 270 – 1.

This discussion suggests that there are at least two categories of cases in which benefitting those other than the risk-imposer overrides the *Nonreciprocity Principle*. The first are those where the victim is the prime beneficiary of the risky activity. Even if the risk-imposer did not act with the explicit authorisation of the victim, she will not be liable for resulting injuries if she acted for his benefit. This is subject to two important qualifications: it does not apply where the benefit is insufficient or if the risk-imposer had a duty to obtain consent before acting for the victim's benefit. Undoubtedly, a surgeon is liable to compensate a patient for operating on him without consent. But where it is not possible to obtain consent, or where there is no duty to seek it, and the prospective benefit justifies the risk, the *Benefit Principle* rather than the *Nonreciprocity Principle* yields plausible results. In the second type of case, the victim is not a beneficiary of the activity, but her claim for compensation against the risk-imposer is outweighed by the benefits of the activity enjoyed by others. If the risk to the non-beneficiary is small and the benefit to others sufficiently great, as with the Amish farmer and the flight path, nonreciprocity will not ground a claim for compensation against the injurer. Note that the *Benefit Principle*, as previously stated, only applies where the victim is the prime beneficiary, and must therefore be amended to accommodate this type of case. Nevertheless, the crucial feature in both types of case is that the distribution of benefits created by the activity outweighs the imposition of a nonreciprocal risk.

5. The Components of Nonreciprocity

So far, I have adduced cases to show that the *Nonreciprocity Principle* is often outweighed by both of the other principles under consideration. Nonreciprocal risk-taking does not ground liability either when the person who is subject to the risk could have avoided it more easily than the risk-imposer or when the person who is subject to the risk is the prime beneficiary (though only when the benefit is sufficient and seeking consent is not possible or

necessary). We can now take a step further and inquire whether nonreciprocity grounds a claim when neither of these conditions obtains. In other words, does the *Nonreciprocity Principle* yield plausible results even when it does *not* conflict with the *Avoidability* or *Benefit* principles?

Consider:

Life Raft: Through no fault or choice of their own, Alice, who has a broken arm, and Rachel are set adrift, in danger of death. Alice can climb aboard a life raft and potentially save their lives, but in order to do so she must clamber over Rachel's unconscious body, risking moderate injury. Alice does this, saving their lives, but in doing so breaks Rachel's arm.

In *Life Raft*, the *Avoidability* and *Benefit Principles* are inapplicable. Both parties lack an adequate opportunity to avoid the risk and the rescue attempt benefits both equally. Conversely, the *Nonreciprocity Principle* applies as Alice risks injury to Rachel but not *vice versa*. We should also assume that Alice already has a broken arm in this case. If not, the inequality of the outcome may give us independent reasons to hold Alice liable to compensate Rachel to some degree.³⁶ Since we wish to isolate the effect of nonreciprocal risk-taking, we should ensure that the outcomes are equal, thus eliminating reasons to transfer resources based on outcome inequality. I will make the same assumption about subsequent variations of *Life Raft*, with the exception of *Life Raft 4*, where my purpose is explicitly to compare the moral relevance of outcomes with *ex ante* risk distribution. With these clarifications made, it still seems that in *Life Raft* Rachel has no convincing claim for compensation against Alice.

³⁶ For a detailed discussion of inequality and the multifarious ways it might be undesirable, see Larry Temkin, *Inequality*, Oxford: Oxford University Press, (1993).

We can further investigate the apparent insignificance of nonreciprocity by evaluating its components rather than comparing it in its entirety to other factors. Recall that nonreciprocity has two components: unequal distribution of risk and asymmetrical creation of risk. I noted in section 1A that these components are distinct. A risk can be distributed equally or unequally between a set of people independently of whether any of them is a cause of the risk. Two drivers parked by the side of the road may be subject to equal risk from a passing lorry, in which case neither of them is a cause of the risk.

We can begin by showing that the asymmetrical causality component of nonreciprocity is insignificant. We can do this by comparing two variations of *Life Raft*. In the first variation, Alice and Rachel will benefit equally from a rescue attempt being performed by Alice *alone*. Alice performs the rescue, imposing an equal risk of harm on both herself and Rachel. As a result, Rachel is injured. In the second variation, Alice and Rachel will benefit equally from a rescue attempt being performed by Alice and Rachel *together*. Alice and Rachel perform the rescue, reciprocally imposing an equal risk of harm on each other. As a result, Rachel is injured.

Let us review the similarities and differences here. In both variations, the risk imposition equally benefits both parties. Both parties are also subject to an equal risk. The outcomes are also identical. The only difference between the two is that asymmetrical causality is present in one but not the other. In the first variation Alice is the only causal source of the risk, while in the second Alice and Rachel both cause risks. But it is still difficult to see how there is a morally relevant difference between the two. Note that we need not come to any decision about whether Rachel ought to be compensated, and if so how much. We need only claim that *if* Rachel ought to be compensated in the first variation then she ought to be compensated in the second, and if she ought not to be compensated in the first then the same judgment should be reached in the second.

What about the second component of nonreciprocity, unequal risk distribution? The distribution of risk does seem to affect liability. In *Life Raft*, we assumed that Alice and

Rachel could only be saved if Alice climbed over Rachel's unconscious body. But consider the following variation:

Life Raft 2: The same as *Life Raft*, except Rachel is not unconscious and either she or Alice can use the other to climb to safety, imposing a risk of moderate injury. Acting unilaterally, Alice uses Rachel's body to climb to safety, saving them both. In doing so, she's breaks Rachel arm.

Is Rachel entitled to any compensation in *Life Raft 2*? Rachel could argue that, although her injury is preferable to death, Alice still did not act fairly. Either of them could have imposed the risk on the other with the same prospect of mutual benefit, and this risk could have been distributed equally by splitting the risk, or if the risk is indivisible, by splitting the odds of undertaking the risk. In other words, unequal risk distribution may be grounds for compensation even when the risk-imposing act is equally avoidable, and for the equal benefit of both parties, *if the ex ante risks were not distributed fairly.*

But although *Life Raft 2* demonstrates the relevance of risk distribution, this does nothing to vindicate the *Nonreciprocity Principle*. This is for two reasons. First, nonreciprocity holds that there may be a case for compensation if the risks are unequal. In describing nonreciprocity as a principle of fairness, Fletcher thus assumes that unequal distributions are unfair. But it is familiar from debates about distributive justice in political theory that this is not a safe assumption to make. And arguments that apply to the fair distribution of resources, welfare, opportunities and so forth, also apply to risk distribution. In some cases, for instance, we may have prioritarian reasons to gives benefits to the worse off rather than distribute them equally.³⁷ This is demonstrated by the following variation of *Life Raft*:

³⁷ On the distinction between priority and equality, see Derek Parfit, "Equality and Priority," *Ratio*, 10, (1997), 202-221.

Life Raft 3: Alice and Rachel are set adrift, in danger of death. Alice and Rachel can distribute the risks of the rescue attempt equally, or one of them can use the other to climb to safety, imposing the risk on them. Rachel has a degenerative disease which means that she generally at risk of losing the use of some of her limbs.

In *Life Raft 3*, Alice ought to accept the risk for the sake of the rescue. Splitting the risk would be unfair in this case as Rachel is already subject to greater risk of physical injury than Alice, and therefore has priority in avoiding further burdens. But the *Nonreciprocity Principle* has the counterintuitive implication that if Rachel performs the rescue by visiting the risk on Alice, as a result of which Alice is injured, Alice will be able to seek compensation from Rachel. By conflating fairness with equality, the *Nonreciprocity Principle* is unable to reach plausible conclusions in this and similar cases.

Second, even if we assume that equal and fair distributions are identical, the relevance of fairness in the distribution of costs goes beyond the boundaries set by the *Nonreciprocity Principle*, sometimes in ways that conflict with the principle. This is because nonreciprocity applies the distribution requirement exclusively to *risk*. But the importance of distribution applies to other costs besides risks, such as *ex post* harms, and sometimes justifies redistribution *despite* initially equal distribution of risk. Keating makes the same point in his defence of strict liability, claiming that “what counts is not reciprocity of *risk* but reciprocity of *harm*... Risk can be fairly distributed while harm is unfairly concentrated, and it is harm that matters.”³⁸

To qualify the argument, it is important to note that this is not true when the *value* of the risk-distributing activity relies on letting wins and losses lie where they fall. If Alice freely gambles her money in a card game with Rachel, she has no claim for compensation against Rachel when she loses her money to her. The value of gambling depends on not

³⁸ Gregory Keating, “A Social Contract Conception of the Tort Law of Accidents”, p. 35.

interfering with the outcome and is in principle a permissible activity (assuming it does not involve severe exploitation or asymmetry in bargaining power). This might be because a freely made decision to gamble reflects a permissible exercise of one's autonomy,³⁹ and perhaps for other reasons. But in other cases we do not have these reasons to let losses lie where they fall. They are not present in the life raft variations we have been considering, where the risks are unwanted but accepted for the sake of mutual benefit. Here, there may be grounds of compensation which are not picked out by the *Nonreciprocity Principle* because the initial distribution of risk is equal:

Life Raft 4: Through no fault or choice of their own, Alice and Rachel are set adrift, in danger of death. They can save themselves by climbing aboard a life raft but each are subject to a risk of serious injury. The attempt will only succeed if they are both involved. They perform the rescue attempt together, imposing equal risks on one another. Alice emerges onto the life raft unscathed but Rachel injures herself and loses the use of her legs.

Does Rachel have a claim for compensation in *Life Raft 4*? The *Nonreciprocity Principle* implies that she does not, since the *ex ante* risks are equal, but this is implausible. As Keating recognises, risks are only significant insofar as they threaten actual harm or loss. We ought to care about the distribution of actual losses more than the distribution of risk. Rachel suffers a significant harm whilst Alice is unscathed. Given that the cooperation of both parties is required to secure the overriding mutual benefit, we have reasons to compensate Rachel to some degree.

³⁹ For a discussion of three different values of choice, see Scanlon, *What We Owe to Each Other*, pp.

251 – 6.

These final arguments show that quite apart from the significance of other considerations that outweigh nonreciprocity, the components of nonreciprocity itself, on closer inspection, do not seem to have much impact on liability.

6. Conclusion

Let us summarise what we have found. Nonreciprocity consists of two distinct ideas; first, that A and B are subject to unequal risk, and second, that A causes a risk to B and not *vice versa*. I sought to identify a convincing moral notion which underpins the *Nonreciprocity Principle*, encompassing both of its components. The claim that A lacks standing to complain against B's risk-imposition because if A has reciprocated the risk then she is a hypocrite covers both components of nonreciprocity. But this suggestion is not convincing for two reasons. First, it fails to explain cases where risks are reciprocated but there is still a claim for compensation, such as the case of the two negligent motorists. Second, it does not apply to permissible activities, yet nonreciprocity seeks to explain liability arising from permissible as well as wrongful acts, so the hypocrisy argument is too limited to support the *Nonreciprocity Principle*.

In order to assess the normative impact of nonreciprocity against other factors such as avoidability and benefit, I then evaluated cases in which the implications of the *Nonreciprocity Principle* conflict with the implications of the *Avoidability* and *Benefit* principles. In both instances nonreciprocity is overridden. In cases where the benefit accrues to the risk-imposer or she could have avoided the risk more easily, nonreciprocity does not ground a claim for liability. We also found that when the other two factors are equalised (when both parties can equally avoid the risk and the activity is to their equal benefit) nonreciprocity still has false implications, as even here there is not always a valid compensatory claim.

Finally, I argued that the *Nonreciprocity Principle* is further undermined if we scrutinise its two components individually. If we vary *Life Raft* to test the significance of asymmetrical causality we find that it is inert. *Life Raft* also showed that risk distribution can affect liability. However, this is because of the independent relevance of distributive fairness. In assuming that fair distributions are identical to equal distributions, nonreciprocity cannot explain cases where one party has a priority-based claim to avoid being burdened with an equal share of the risk. Furthermore, in applying the criterion of equal distribution exclusively to risk, nonreciprocity fails to take into account the effect of distributive principles on *ex post* as well as *ex ante* burdens. *Life Raft 4* suggested that we sometimes have reasons to redistribute outcome costs even if *ex ante* risks were distributed fairly. In this way, the independence of distributive fairness can count *against* the *Nonreciprocity Principle*.

For these reasons, we should reject the claim that nonreciprocal risk-taking is the moral basis of liability to compensate. The results of this critique are not purely negative, however. In outlining and assessing the *Avoidability* and *Benefit* principles, we have made some progress towards a more satisfactory account of liability to compensate. The *Avoidability Principle* captures the insight that a person can be expected to bear a greater burden if she had a good opportunity to avoid the risk. This insight already underlines some defences in tort, such as contributory negligence, but the *Avoidability Principle* plausibly holds that the injurer's opportunity to avoid imposing the risk is equally relevant.

The *Benefit Principle* is distinct from the *Nonreciprocity Principle*, but the two are sometimes conflated. The assumption that individuals who impose reciprocal risks benefit equally is false, but may explain why the *Nonreciprocity Principle* looks initially appealing. The distribution of benefits produced by a risk-imposing activity is clearly morally relevant and different activities can yield a complex variety of benefit distributions, a full investigation of which is beyond the scope of this chapter. It is also important to focus on the *actual* outcomes of risky activities, both beneficial and harmful. Considerations of distributive fairness are relevant, but should be applied to the distribution of harms and

benefits as well as *ex ante* risks. Establishing that these factors are important is not to formulate a full account of the moral basis of liability to compensate, but it is a step in that direction.

Corrective and Distributive Justice

In chapter 6 we investigated the moral basis of corrective duties, where I claimed that considerations of distributive fairness are relevant in deciding who should bear the cost when one person harms another. This claim indicates a closer relationship between corrective and distributive justice (hereafter CJ and DJ) than many have assumed. The distinction between these two forms of justice has its roots in Aristotle, but has recently garnered much scholarly attention. The first aim of this chapter is to reject the claim of some theorists that CJ is independent of DJ. This is really a set of different claims, as there are a number of ways that corrective and distributive justice might be independent of each other. I distinguish between these claims and reject each of them. Showing that CJ is not independent of DJ will justify a presupposition that was necessary for my previous claim, in chapter 6, about the relevance of DJ for corrective duties.

Throughout this thesis, I have also employed the language of *corrective* duties. This usage might suggest that corrective duties are a deontological category of fundamental interest, and that this category is of special relevance to tort theory. The second aim of the chapter is to determine whether this suggestion is true. In order to do this, we will define corrective duties more carefully and compare the distinction between corrective and non-corrective duties with another between agent-centred and agent-neutral duties. We will find that despite the language commonly employed, including the terminology in earlier sections of this thesis, the latter distinction more accurately frames the relevant issues for tort theory.

1. What is Corrective Justice?

Numerous theories of private law can be grouped under the rubric ‘corrective justice’. Theories that fall under this banner are diverse and it is not easy to identify which features

they have in common, in virtue of which they are theories of CJ. One claim they tend share is that private law embodies CJ, particularly in its structural features. For example, it is concerned with the adjudication of claims brought by one private party against another. In such cases a court may, among other things, decline to offer the victim a remedy, impose full compensatory damages on the defendant, or apportion responsibility between the parties. Although the court is therefore concerned with an allocative question – how to distribute *ex poste* costs between two parties – this differs from the allocative questions that occupy theorists of DJ. One difference is that DJ is concerned with allocative issues across larger groups, such as citizens of a political community or human beings *tout court*.

This is not to say that CJ is distinguished by the size of the group that brings or responds to the civil action. Corporations and governments can be sued and classes can bring legal actions as well as individuals. Nevertheless, a civil action links two parties and no others, even if ‘party’ is understood to include entities consisting of more than one individual. This direct connection between a particular plaintiff and a particular defendant is sometimes called the ‘bipolarity’ of civil suits. As Ernest Weinrib has put it, this ‘bipolarity’ is the “master feature” of private law.¹

Does this feature help us distinguish corrective from distributive justice? We could understand the feature in a non-institutional and non-procedural way, as embodying the following question: between two given parties, one of whom is harmed, who should bear the *ex poste* cost of the harm? This definition is very general. It does not require that one party commits a wrong against another, or that the wrong causes the harm. It follows that a decision about which of two people living on a desert island should bear the cost of an injury caused by a falling coconut is a matter of CJ. This seems wrong. Indeed, it looks like a distributive matter involving only two individuals. The mere fact that an allocative question

¹ Ernest Weinrib, *The Idea of Private Law*, Cambridge, MASS: Harvard University Press, (1995) at p. 10.

involves two parties (even when a party can be a corporate entity) is not sufficient to bring it within the scope of CJ.

There are two other common ways of caching out the ‘bipolarity’ feature: appealing to causation and the breach of a directed duty. A directed duty is a duty owed to another individual or group.² The breach of this duty is therefore a wrong *to* that party. Duties in private law are generally of this kind, such as duties not to commit batteries, and they are distinct from public law duties that are not owed to anyone in particular, such as duties not to destroy the breeding sight of a protected species. The second way to understand bipolarity is that one agent must cause harm to another. Causation provides a straightforward link between two parties to a private action.

Weinrib emphasizes the first feature but rejects the idea that mere causation is sufficient.³ Private law is a law of directed wrongs. Liability founded on mere causation is alien to corrective justice because it does not involve wrongful conduct. Others such as Coleman accept that strict liability can be a matter of CJ.⁴ Coleman does this by labelling the *loss* wrongful, even though it is brought about by a permissible act. The loss is wrongful because it follows from the infringement of a right. The claim generated by this right does not disappear simply because the defendant acts justifiably.⁵

² See chapter 1, section 2B.

³ *Ibid.* at. P. 177 – 9.

⁴ Coleman, *Risks and Wrongs*, pp. 371 – 2.

⁵ He applies this model to the rule in *Vincent v Lake Erie* and Joel Feinberg’s cabin case. *Vincent v Lake Erie* held that a person acting for reasons of necessity is liable for damage he causes to the property of others. In Feinberg’s cabin case, a hiker brakes into a cabin to avoid freezing to death. Although he acts justifiably, he still owes the cabin owner compensation. See Joel Feinberg, “Voluntary Euthanasia and the Inalienable Right to Life”, *Philosophy and Public Affairs*, 7, (1978), 93 – 123, p. 102.

From the perspective of tort practice, it is questionable whether either of these features alone defines CJ. Strict liability is usually considered part of tort law, and actions in strict liability seem as much to do with CJ as fault-based torts. Moreover, proof of causation is not always necessary. Situations involving insurmountable evidential difficulty have given rise to well-known exceptions to the requirement to prove causation.⁶ We could adopt a disjunctive account and hold that strict liability and exceptions to the usual rules of causation are both compatible with the bipolarity of private claims. CJ involves the allocation of *ex poste* costs in situations where *either* one agent causes harm to another *or* one agent wrongs another (or both). If each of these features entails the bipolarity of a civil claim, it is not unexpected that they should both be components of CJ.

Given the controversy mentioned in chapter 1 about the proper boundaries of tort law, perhaps it is better to adopt a more philosophical approach to carving up the conceptual territory. We should ask which concepts and distinctions best track matters of moral interest, rather than currently valid legal categories. Let us therefore take a few steps back and build up a taxonomy of the ways that CJ might be understood. We will start at the most general level and proceed by introducing narrower versions. We will then be in a position to contrast this family of concepts with DJ. Each account of CJ should answer at least three questions: (1) what actions count as corrective? (2) What is corrected? (3) Who does the correcting? For now, we can omit an exhaustive list of actions that count as corrective. I have discussed this issue previously, although not exhaustively,⁷ and the relationship between corrective and distributive justice does not depend on resolving it. From here on, I will refer to an unspecified list of such actions as ‘corrective actions’.

⁶ The most prominent example in English law is compensation for mesothelioma. See especially *Fairchild v Glenhaven Funeral Services Ltd* [2002] UKHL 22 and *Sienkiewicz v Grief (UK) Ltd* [2011] UKSC 10.

⁷ See chapter 4.

On any account, CJ must involve correcting something, but there are many ways to answer question (2). Here are a few. The most general can be called *The Harm View*.⁸ This holds that CJ consists in correcting harms. I understand harm as a setback to a good or goods enjoyed by some individual. Previously, I argued that we should think of harm more specifically as setback to *wellbeing* (rather than resources or preference satisfaction),⁹ but that problem need not concern us here. The important feature of this view is that it encompasses harms that were not brought about by human agents, such as injuries caused by falling rocks or bolts of lightning.

Some theorists take it as self-evident that this is not the right way to understand CJ. Coleman and Weinrib both implicitly reject it, and Stephen Perry refutes the view that CJ fixes deviations from distributively just patterns on the grounds that it entails what I have called *The Harm View* (after all, natural events also disrupt distributive patterns).¹⁰ In other words, he assumes that entailing *The Harm View* is a valid *reductio* argument for a thesis about CJ. If we are constrained to interpret CJ in light of legal practice, this assumption is correct. But from a purely normative perspective, whether they are right to dismiss this view depends on whether there are moral reasons to treat harm caused by human actions differently from harm caused by natural events. I will return to this in section 2.

⁸ Perhaps the view should be broader still and include correcting detrimental states other than harms. This might include removing threats or risks, which are not harmful in themselves, or symbolic corrective actions such as apologies. I assumed in chapter 4 that removing a threat was a corrective action, but this might be questioned. This serves to highlight the conclusion drawn there that what matters is that the duty to remove a threat is *preventive*, not necessarily that it is corrective.

⁹ See chapter 3.

¹⁰ Stephen Perry, "On the Relationship between Corrective and Distributive Justice", in *Oxford Essays in Jurisprudence*, Ed., Jeremy Horder, 4th series, Oxford: Oxford University Press, (2000), p. 237.

Despite its broad scope, *The Harm View* excludes cases where people are put into bad states but are not made worse off. Being made worse off is a comparative notion: it indicates that one is worse off compared to some relevant alternative. Perhaps one is worse off than one was before some event (the temporal view) or worse off than one would now be had the event not occurred (the simple counterfactual view). But not all bad states depend on this comparative idea. Some states are non-comparatively bad. Consider a person born with a congenital disability. This person is not worse off than *she* would be in some alternative state because there is no alternative in which *she* exists. The events that led to her disability are her conception and birth. She is not worse off than she was before these events because she did not exist before them. Equally, she is not worse off than she would now be if not for these events because she would not exist if not for them.

There are two more answers to question (2). One is *The Human Agency View*, which holds that CJ consists in correcting only harms caused by human actions. *The Human Wrongs View* is narrower still and singles out harm caused by wrongful conduct. The common assumption on both of these views is that CJ is only concerned with harm caused to humans. But it is hard to see why. Non-human animals can be harmed both wrongfully and in the course of permissible activity. Of course, they cannot be party to a private action, but the concept of CJ is clearly broader than the way it features in legal practice. If we defend, say, *The Human Wrongs View*, and argue that we have particular reasons to care about harm caused by wrongful acts, we need some further reason to exclude non-human animals that are wrongfully harmed.

It is important to note, at this point, that these views are silent about issue (3): who should do the correcting. With respect to harm caused by human agents, it is natural to assume that the harm-causing agents should do the correcting. Coleman and Weinrib's emphasis on the relational nature of CJ suggests that this is the crucial point. The breach of a directed duty resulting in harm is the standard reason why we should identify certain individuals as bearing a remedial duty. But this is not the only available view. We could instead hold that harm should be corrected by a group of individuals who have committed

the same wrong, regardless of whether it caused harm to the claimant. On this view, negligent drivers would collectively compensate victims of negligence on the road. A related view is that those who create risks should collectively compensate for harm caused when those risks manifest. This expands the group of potential defendants from wrongdoers to risk-creators. Alternatively, society at large could bear the entire cost of compensation, or could share the cost with those more directly involved in the harm. I will not attempt to evaluate these views here. This brief treatment is merely meant to illustrate that there are many possible positions one might take on the three questions that any theory of CJ must answer.

2. Corrective and Non-Corrective, Agent-Centred and Agent-Neutral

How should we delimit the scope of corrective duties? To begin, let us note a crucial restriction on corrective duties: responses to non-comparatively bad states cannot be corrective. This might be thought surprising as it excludes a category of cases that form part of tort law, namely ‘wrongful life’ cases. But correction seems to be an implicitly comparative idea. Correcting for some harmful event presupposes that the event has reduced a person’s wellbeing, resources or some other good. This is not true when a person would not have existed had the event not occurred.

Perhaps we could conceptualise responses to these bad states as corrective in some other way. We might say that they place a person below a defined standard of wellbeing, such as that required for a minimally decent life. People who would not have existed if not for the event that caused their disadvantage may still be worse off than an objective standard. The problem with this approach is that it renders talk of correction redundant. If the important issue is whether people are located below some threshold of wellbeing, it would be better to dispense with talk of CJ altogether and focus only on DJ. A duty to benefit a person born with a congenital disability is therefore not a corrective duty, even if

the disability is causally dependent on wrongful conduct. But if wrongful life cases are morally and legally important, do corrective duties represent a category worthy of investigation in its own right?

To make progress with this question, let us introduce some distinctions. The *agent-relative* perspective is that of an individuated agent. This perspective is important when reasons or duties apply to particular agents and not others. Agent-relative duties may arise when an agent wrongs or harms another, as these events create a special relationship between the agent and her victim that the agent does not have with others. The *agent-neutral* perspective is that of any agent, including those who are not linked to the victim. When there are no facts that give rise to special reasons or duties towards a victim, a person's agent-relative reasons are also her agent-neutral reasons.

We have already encountered the next relevant distinction. It is between harm (which is implicitly comparative) and non-comparatively bad states. Since correction, like harm, is also implicitly comparative, one cannot correct non-comparatively bad states. Harms can be further subdivided into harm caused by natural events (natural harms), permissible human action (agential harms) and wrongful conduct (wrongful harms). Altogether, then, there are at least four categories of interest: non-comparatively bad states, natural harms, agential harms and wrongful harms. There may agent-relative and/or agent-neutral duties and reasons produced by all four of these states. Corrective duties are a significant category *only if* there are reasons or duties produced by the three types of harms that are not produced by otherwise equivalent non-comparatively bad states.

We can test whether there is a difference between harm and non-comparatively bad states from both an agent-relative and an agent-neutral perspective. Let us begin with the agent-neutral perspective. We can compare four cases that are equivalent except that they involve a non-comparatively bad state, a natural harm, an agential harm and a wrongful harm respectively. Here are a set of cases that meet this criterion:

Egg Selection: An egg is selected for a couple undertaking a programme of assisted reproduction. Unfortunately, the egg has a defect which means that the child, Davesh, is born with a genetic disorder. This disorder is detected from early age, but there is no cure. It causes Davesh to lose the ability to walk when he is 8 years old.

Fallen Tree: Richard is 8 years old. A tree unexpectedly falls on his leg, removing his ability to walk.

Motorbike Accident: A woman is riding her motorbike. Although she pays due care and attention to her surroundings, she runs into Miles, who is 8 years old. As a result of his injuries, Miles loses his ability to walk.

Reckless Biker: A drunken woman is driving her motorbike recklessly and runs into John, who is 8 years old. As a result, John loses his ability to walk.

From the agent-neutral perspective, do we have reason to favour any of these children over the others? Suppose we could only restore the mobility of one child, should the means by which he lost his ability to walk make any difference to our decision? I doubt that it does. The appropriate solution is to give each child an equal chance of being benefitted.

Victor Tadros argues that we have no reason to benefit victims of injustice over those who are similarly harmed by natural events. He imagines two groups of children who have been infected with HIV. One group has been infected as a result of bad luck, while the other has become infected due to the government's unjust health policies.¹¹ When choosing which of these children to benefit with antiviral drugs, we have no reason to prefer one group to the other. Indeed, it would be wrong to place a child at the top of the queue because his infection was caused by injustice. If wrongful harms do not give us agent-neutral reasons

¹¹ See Victor Tadros, "The Ends of Harm", p. 106.

that natural harms do not, then the same must be true of agential harms. This means that we have no reason to prefer Richard, Miles or John when deciding who to benefit.

The important question, however, is whether we have reason to prefer these three to Davesh, as Davesh has not been harmed – he is in a non-comparatively bad state. One reason is that Richard, Miles and John have suffered some ancillary misfortunes that Davesh has not. First, because they were not expecting to be harmed, they have already formed plans and set ends that are mobility dependent, and these plans will be frustrated by the loss of their legs. Davesh’s plans have not been frustrated because he has known of his disorder from an early age and has had the opportunity to plan around it. This difference, however, is only contingently related to comparative and non-comparatively bad states.

We can adjust the cases to demonstrate this. Imagine that Richard, Miles and John were part of a health initiative whereby babies were implanted with microchips to prevent a battery of common childhood diseases. Once the chips were implanted, they could not be removed without killing the patients. The experiment went wrong, and it was discovered that the chips would malfunction after 8 years, causing the children to lose the use of their legs. In this variation, the harm is not unexpected and the victims are able to plan around it. But Richard, Miles and John are still comparatively worse off. They are worse off because of the experiment than they would now be had it not taken place, while Davesh is not worse off at all. He is still better off with his injury because he would not exist had it not occurred.

There are other contingent factors that may legitimately affect our decision making. Real cases of non-comparative disadvantage often involve birth defects. This means that these victims do not suffer *transitional* harm. A transitional harm is a harm that is suffered by the process of having a reduction in one’s wellbeing, and may include feelings of loss and painful memories. As Dante wrote, “There is no greater sorrow/Than to be mindful of the happy time/In misery”.¹² On the other hand, those who suffer transitional harms have lived a portion of their lives with whatever good they have lost, while those born with birth defects

¹² Dante Alighieri, *The Inferno, Canto V, Second Circle*, Translated by Henry Wadsworth Longfellow.

have never experienced the same privilege. As Alfred Lord Tennyson wrote, “’Tis better to have loved and lost/Than never to have loved at all.”¹³ These considerations are relevant, but as before they are only contingently related to the distinction between comparative and non-comparatively bad states.

Ultimately, it is doubtful that there is any non-contingent reason to benefit those who are harmed over those who are in non-comparatively bad states. Even the contingent reasons may have little relevance. In the four cases above, each patient should be given an equal chance of receiving the benefit even if Davesh is able to plan around his injury. The ability to plan is not enough of an advantage to systematically prefer those who are harmed over those who are in equivalent states but who have not been made worse off. From the agent-neutral perspective, what matters is how well off people are, not that they are worse off compared to some alternative state.¹⁴

What about the agent-relative perspective? Agent-relative factors may establish whether a duty exists at all, or the stringency of that duty. The stringency of a duty specifies the cost that an agent must incur in order to fulfil it. The more stringent the duty, the greater the cost the duty-bearer is required to shoulder for the sake of the duty. It is uncontroversial that people who wrongfully harm others bear more stringent duties towards their victims than bystanders. This is true from both the *ex ante* and *ex poste* viewpoint. If A maliciously roles a boulder down a hill towards a victim, he has a more stringent duty to avert the threat than a bystander.¹⁵ If the only way to save the victim’s life is to give up his own, he is required to do so, while a bystander is not duty-bound to make such a sacrifice to save the victim. Once the boulder has caused an injury, the wrongdoer also has a more stringent duty to compensate the victim than a bystander.

¹³ Alfred Lord Tennyson, *In Memoriam, Part XXVII*, (1850).

¹⁴ Derek Parfit discusses a related view about whether we have special reasons not to harm a person compared with creating a badly off person. See *Reasons and Persons*, pp. 366 – 371.

¹⁵ See Victor Tadros, *The Ends of Harm*, especially chapters 11 and 12.

It is more controversial whether those who cause harm through permissible actions have more stringent duties than bystanders. Some hold that only wrongdoing can create a more stringent duty. Others hold that harming people through permissible actions can create a more stringent duty, but only if the harm was foreseeable. Others hold that merely causing harm can create a more stringent duty, even if the harm was not foreseeable. For present purposes, we need not entangle ourselves in the significance of causation, foreseeability and wrongdoing. Let us stipulate that all of them make a difference from the agent-relative perspective by increasing the stringency of a person's duty to avert a threat and compensate for an injury.

If this stipulation is true, it would make a great difference to normative tort theory, for which the significance of wrongdoing, foreseeability and causation on the duty to compensate are central concerns. But none of these three factors are limited to corrective duties. To see this, compare *Reckless Biker* with a variation of *Egg Selection*. In this version, the doctor recklessly selects an egg without screening it, which leads to Davesh's condition. Like the negligent biker, the doctor acts wrongly and causes Davesh's disability, but unlike the biker the doctor has not made Davesh worse off and so cannot owe him a corrective duty.

I will not attempt to determine whether the positions of the biker and the doctor are *equivalent*. But it seems that the presence of wrongdoing and causation in *Egg Selection* has *some* effect on the doctor's relationship towards Davesh, which distinguishes him from a bystander. Perhaps the doctor has an agent-relative duty to benefit Davesh because of his wrongdoing. As long as this modest claim is true – that wrongdoing, foreseeability and causation make some difference in non-correction cases – corrective duties are not a category of fundamental interest.

This is because these factors are not related to harm. *Egg Selection* shows that they can be present when no harm is caused to anyone. Instead, they are captured with the broader concept of agent-relative duties. The duties we have to benefit others depend on our relationship to them. If we wrong others or cause them to be in bad states, this may alter that

relationship by creating more stringent duties to benefit them. These factors are only contingently related to harming and correcting, but not so with agent-relativity. The concepts of ‘cause’ and ‘directed wrong’ are necessarily relational. A person causes *something* and a directed wrong is, by definition, a wrong to *someone*. This means that they affect the relationship between an agent and the person he wrongs or causes to be in certain states. Fundamentally, then, it is the category of agent-relative duties rather than corrective duties that captures the significance of wrongdoing, foreseeability and causation to our duties to benefit others.

3. The Relationship between Distributive and Corrective Justice

Many defenders of CJ have claimed that principles of liability in private law are independent of public or distributive goals. Weinrib argues that private law is both conceptually and normatively independent of distributive goals. It is conceptually independent because relying on distributive considerations, such as the use of insurance to spread losses, introduces incoherence into tort law. It does so because the loss spreading rationale gives us no reason to single out a particular defendant as the subject of a legal action – the rationale would apply against any member of the insurance pool.¹⁶ In short, it cannot account for bipolarity, the ‘master feature’ of private law. It is normatively independent for similar reasons. Weinrib claims that “the private law relationship... is a mode of moral association that attaches decisive importance to the justification of the norms that constitute it. Coherence is important for private law relationships because it is indispensable to justification”.¹⁷

These remarks leave many questions about the relationship between corrective and distributive justice unanswered. In what follows I will outline a number of ways in which

¹⁶ Ibid. at p.36.

¹⁷ Ibid. at p. 39.

this separation between the two types of justice cannot be maintained. Distributive concerns always seem relevant to the third question that any theory of CJ must answer: who should do the correcting? Perhaps, as Weinrib insists, it is only those who violate the rights of others who should do the correcting. But if this is a normative claim it requires a defence. We need to know why the wrongful violation of a right outweighs distributive concerns in favour of sharing the costs of accidents. As Jeremy Waldron has emphasised, some moments of carelessness are innocuous while others lead to massive loss.¹⁸ If there is a high level of outcome inequality between negligent actors, surely distributional issues must be *relevant* to CJ. This is true even if they are only relevant insofar as they are overridden by agent-relative reasons to impose corrective duties solely on injurers. CJ theorists have generally failed to address these points. Weinrib, for example, excludes distributional issues by fiat, on the grounds that their consideration would render private law incoherent. But even if we accept Weinrib's dubious claim that the coherence of private law plays a key role in its justification, the value of this coherence itself must be justified against countervailing arguments that private law should be consistent with DJ.

Similarly, it may not be permissible for legal institutions to participate in distributive injustice. The permissibility of implementing CJ may be conditional on the social distribution of goods meeting basic standards. It is certainly implausible that implementing CJ is permissible *regardless* of how resources in society are initially distributed. If resources are distributed between two people in a drastically unequal fashion, so that one is wealthy whilst the other is impecunious, would it be justifiable to hold the poor person liable to fully compensate the rich person, rendering her destitute in the process, if she negligently damages the rich person's Ferrari?

Other writers who accept the independence thesis in some form concede this point. Stephen Perry, although he rejects the view that the function of CJ is to preserve determinate

¹⁸ Jeremy Waldron, "Moments of Carelessness and Massive Loss" in *Philosophical Foundations of Tort Law*, Ed., David Owen, Oxford: Oxford University Press, (1995).

just distributions, does not insist that enforcing corrective justice is permissible irrespective of the prior allocation of wealth. We can agree with Perry that CJ does not exist merely to maintain patterns of wealth, akin to redistributive taxation. As Perry argues, such a view fails to capture two things. The first is the different justificatory bases of claims in CJ and DJ. The moral focus of the victim's claim is the harm she has suffered at the hands of another responsible agent, rather than their different level of resources. Second, claims in CJ are not complaints about systemic patterns such as widespread social inequality, but individual cases. These points can be accepted, but as Perry recognises, CJ and DJ are still sources of moral claims that must be balanced against one another.¹⁹ To see the importance of this point, suppose that there is a causal connection between the two forms of wrongs, such that distributive injustice greatly increases instances of negligence. It is hard to see why this fact would have no relevance to the justifiability of negligence liability schemes.²⁰

Thirdly and relatedly, distributive concerns might place limits on the enforcement of CJ. Assuming a commitment to treating like cases equally, according to the CJ model, if two people commit an identical wrong, causing identical harm, they both ought to be held liable to the same degree. However, this burden of liability may differ in the effect it has on the wellbeing of each defendant. If one defendant would be rendered almost destitute as a result of fulfilling her duty to compensate her victim, it may be impermissible to enforce that liability. We have already encountered considerations of this sort in chapter 5.²¹ This limit is plausibly a matter of DJ. Enforcing a person's liability may place her in a position that is unacceptable from a distributive point of view. One way to put this is that liability in private

¹⁹ See Stephen Perry, "On the Relationship between Corrective and Distributive Justice", p. 259.

²⁰ For an argument that the conditions of poverty that increase the likelihood of criminal offending can undermine the state's standing to prosecute offenders, see Victor Tadros, "Poverty and Criminal Responsibility", *The Journal of Value Inquiry*, 43 (3), (2009), 391 - 413

²¹ See chapter 5, section 3.

law – and liability to suffer a burden generally – does not imply that it is always permissible to enforce that liability.

Fourthly, the practice of CJ may *exacerbate* distributive injustice. Tsaachi Keren-Paz argues that tort law has a regressive bias in a number of ways.²² The principle of *restitutio ad integrum*, which holds that compensation must place the victim in the position she would have been in had the wrong not occurred, means that it is cheaper to impose risks on the poor, since they can be compensated at lower cost. If this is correct, tort law creates a perverse incentive to prefer risks to the poor rather than the wealthy that re-entrenches existing economic inequalities. Moreover, when offering liability insurance, insurers will collect a premium based on the loss of the average victim. Insured groups pay equal premiums but victims collect unequal compensation depending on their pre-accident wealth. As a result, the poor cross-subsidise the wealthy in the pricing of liability insurance.

Why do Weinrib's claims about the normative independence of private law seem untenable in the face of these criticisms? The fundamental reason is this. According to most principles of DJ (with the possible exception of 'unpatterned' libertarian theories) certain distributions are unfair. These unfair distributions confer reasons to effect a fairer allocation through redistributive mechanisms. These are general *pro tanto* reasons. Unfair distributions, wherever they are found, always give us reasons to redistribute (though they do not always give us all things considered *duties* to redistribute). In this respect, reasons for action given by unfair distributions are similar to other moral reasons, such as those given by the suffering of others. The reasons to alleviate or prevent suffering do not disappear because suffering is produced by an institution with its own moral practices; they are given by suffering *period* and any practice that causes or even fails to prevent suffering may have to justify itself against these *pro tanto* reasons. As a rough analogy, imagine a harsh music instructor whose method of teaching and discipline effectively instils musical talent but causes much suffering as a by-product. Such a method could not be justified simply by

²² Tsaachi Keren-Paz, *Torts, Egalitarianism and Distributive Justice*, Ashgate, (2007), pp. 67 – 9.

pointing out that it is internally coherent and the suffering it produces forms no part of its conceptual structure. Such a method could only be justified by showing, at the very least, that its benefits override the suffering it causes.

4. Does Corrective Justice Necessarily Raise Issues of Distributive Justice?

Another way to undermine the independence of CJ is to argue that its operation necessarily raises distributive questions. John Gardner advances such an argument.²³ He begins by claiming that institutions of CJ may still be required in a distributively just society. In this utopia of socio-economic justice, people still negligently run into each other's cars, spread false rumours, break contracts and so on. Unless one implausibly conceptualises *all* forms of wrongdoing as allocative injustices, in which case they would also be excluded *ex hypothesi* from a distributively just society, some wrongs will persist and may require legal institutions to force perpetrators to provide remedies to victims.

Gardner then argues that when we wonder how to organise, subsidise or facilitate the doing of CJ we unavoidably encounter distributive issues. If this is correct, his argument identifies a type of local DJ which is peculiar to the doing of CJ. Even if society is distributively just in all other respects, the introduction of private law institutions will raise new distributive questions.

In brief, Gardner's argument is as follows. The classification of a wrong as a 'tort' gives victims of the wrong a complex legal right that places at her disposal the public authority of the courts (we should understand the good to be distributed as the opportunity to vindicate a right rather than the right itself. The mere possession of a formal right, without the ability to exercise it, is of more limited interest). The victim is given a right, not only

²³ John Gardner, "What is Tort Law For? Part 2", (forthcoming).

against the wrongdoer, but also against the courts, to assist her in her attempt to win a remedy from the defendant. The question of *which* wrongs should be classified as torts raises questions of DJ because the law is “selecting some people for a measure of official support in their personal affairs which most other clients of the welfare state can only dream of.”²⁴ To illustrate his view, Gardner gives a couple of examples from English tort law. First, the granting of tort-law rights to trespassers injured by negligent occupiers of land can be justified because “the fact that trespassers are tortfeasors has been disproportionately visited upon them by too often denying them causes of action in tort for wrongs committed against them by occupiers.”²⁵ Similarly, liability in negligence under the rule in *Donoghue v Stevenson* reflects the fact that retail consumers had previously been unfairly treated relative to those enjoying privity of contract with manufacturers and wholesalers. This unfairness, Gardner claims, is distributive in nature: “plaintiffs of a certain class have been given undeservedly little in the distribution of causes of action in tort”.²⁶

The conclusion of Gardner’s argument seems correct, but it is not clear that these examples succeed in making the point. Gardner is suggesting that, the reason duties of care in negligence are afforded to trespassers is because trespassers as a group are subject to distributive unfairness. One way to interpret this claim is comparative: their position is unfair relative to other groups who are subject to the negligence of occupiers, such as invitees, who are afforded decent rights. The categorical denial of rights against negligent injury to one group compared with the standard rights afforded to the other violates principles of distributive fairness. There are a few qualifications that must be added to this view.

First, the conferral of these tort rights is defensible in some cases that do not involve distributive unfairness, so these changes in the allocation of basic rights do not entirely

²⁴ Ibid. at p. 14.

²⁵ Ibid. at p. 16.

²⁶ Ibid.

depend on arguments from DJ. If occupiers owe duties to trespassers when invitees do not enjoy similar duties, the justificatory work cannot be done only by the unfair position of trespassers relative to invitees, because both are equally denied such rights. Suppose we lived in a society in which no duties of care were owed by occupiers at all. In such a society, rights against negligent injury would not be distributed unfairly between trespassers and invitees, since neither group would have any. We would have no reasons of DJ, of the kind identified by Gardner, to recognise duties to trespassers. But it is plausible that we ought to recognise such duties regardless. There are a number of considerations that tell in favour of this conclusion that have nothing to do with distributive concerns.

To illustrate these, let us consider a hypothetical case:

Negligent Occupier: Duties of care in negligence are not afforded to trespassers or invitees. Nevertheless, trespassers are injured by negligent occupiers far more frequently than invitees, for a number of reasons. Most dangerous features are outside buildings; trespassers, unlike invitees, are never warned of the danger; and there is no incentive to reduce the risk of injuries. Peter accidentally trespasses on Stephen's land, which is dangerous and poorly signposted. Stephen is aware of the danger, knows that accidental trespassers regularly enter the vicinity of the danger, and could take precautions at minimal cost. However, he doesn't bother to take precautions and Stephen is seriously injured.

In *Negligent Occupier*, the difficulty for Peter in avoiding the risk and the ease with which Stephen could remove the danger support the moral case for holding Stephen liable to compensate Peter. This is true despite the fact that invitees are not owed duties of care in negligence and there is no unfairness between the two groups with respect to the distribution of tort rights. Perhaps there is a distributive injustice in the broad sense that a particular good – a tort right – is not being distributed to anyone. But this amounts to saying that invitees

and trespassers sometimes have justified claims that are not legally recognised. If this is a distributive issue, it is not the kind that seems to interest Gardner.

Secondly, Gardner does not focus solely on unfairness between two groups. He claims that some classes of potential claimants have undeservedly little in the way of tort rights full stop.²⁷ This, for Gardner, is still a distributive concern. So he could say that, in *Negligent Occupier*, Peter has undeservedly little full stop and this is a matter of DJ. But Gardner does not explain what is meant by desert in this context. Presumably the reference to desert is not just a way of summing up the various features of the case that justify a claim to compensation. In that case, it is only trivially true that Peter has ‘undeservedly little’. This move would also make Gardner guilty of the error he exposes in those who attempt to conceptualise all wrongdoing as a form of DJ, for it is difficult to see how the features of the case mentioned above that justify liability have anything to do with desert. Gardner must either provide a convincing account of desert or abandon the non-comparative element of the argument.²⁸

Finally, Gardner’s reference to the comparative positions of groups should be read as a mere useful generalisation. Groups are not the fundamental objects of DJ, individuals are. The comparative unfairness between trespassers and invitees would not be a distributive issue if everyone had the same chance of being a trespasser or an invitee. If that was the case, no individual would be worse off relative to any other, even though at the group level trespassers would still be worse off relative to invitees. Many trespassers with justified claims would still be denied a right of action, but comparative unfairness would cease to be a concern. The distributive argument depends on the idea that the likelihood of any individual being in either group is unequal.

²⁷ Ibid.

²⁸ Gardner elaborates on desert, in the context of self defence, in other work. See John Gardner and François Tanguay-Renaud, “Desert and Avoidability in Self-Defense”, *Ethics*, 122(1), (2011), 111 – 134.

With these qualifications made, Gardner's basic position is correct. The opportunity to vindicate a tort right is a useful and fundamental resource, and distribution of these goods cannot be ignored. For example, imagine that in *Negligent Occupier*, duties of care are owed to invitees but not trespassers. Assume also that the poor are much more likely to be trespassers than invitees. Finally, independent reasons to afford tort rights to trespassers and invitees, in carefully defined cases, are equal. In other words, both predefined groups of claimants have equally justified claims to compensation. In this variation there is an *extra* reason, along with those noted in *Negligent Occupier*, to recognise Peter's claim: he is a member of a group that is comparatively disadvantaged with respect to such claims.

5. The Relevance of Distributive Arguments to Standards of Liability

A common criticism of CJ is that it is largely formal.²⁹ It holds that agents have duties to correct wrongs, or wrongful harms, but does not offer a substantive account of the standards of liability that courts should use to determine when an agent bears a corrective duty. Some CJ theorists such as Coleman and Weinrib give some thought to the appropriate standards of liability to supplement the formal principle of CJ.³⁰ Unfortunately, these comments are incomplete and a full account of the basis of liability is still lacking.

How should we decide which standards of liability to adopt? I made some headway on this question in chapter 6, where I suggested that the *Avoidability Principle* and the *Benefit Principle* go some way to creating a full account of the moral basis of corrective duties. I also suggested that distributive considerations relevant. The failure of CJ theorists to produce a complete account means that this cannot be ruled out.

²⁹ See Peter Cane, "Corrective Justice and Correlativity in Private Law", *Oxford Journal of Legal Studies*, 16(3), (1996), 471 – 488.

³⁰ See introduction, part 1C.

To investigate this possibility further, we can begin by distinguishing between two sorts of cases: those in which loss spreading is possible and those in which it is not. It is true, of course, that realistically loss spreading is always possible, since we are interested in the construction of social institutions to govern the costs of accidents in large societies. But it does not follow that there are no moral answers to these questions when losses cannot be spread. First we ought to determine which principles are most defensible in two party cases, before asking how these principles are modified when loss spreading is an option. To focus on the moral considerations that operate in these cases, we can imagine them taking place on a desert island rather than a populous society. This device is designed to exclude the possibility of loss spreading, forcing us to focus solely on the two parties involved. We can then compare our judgements in these cases with equivalent cases where loss spreading is possible. If we find that loss spreading makes a difference to our judgments about what standards of liability we ought to adopt, this will show that distributive factors are relevant.

We should be aware of a number of complexities related to the role of loss spreading in the argument. First, the possibility of loss spreading might determine two things: whether a victim has a claim to compensation and who should satisfy it. Suppose A and B are on a desert island and A harms B in a way that would ordinarily make A liable to compensate B. As we have already seen, whether B's claim for compensation is enforceable depends on the state of affairs that would result from A compensating B. Suppose further that B enjoys a decent level of wellbeing, despite having been harmed by A, and if A compensates B then A would render himself very badly off. Under these circumstances, perhaps it is not permissible for B to force a claim against A. B is not too badly off and A would be rendered destitute by enforcing his liability to B. But if we transpose the case to a large society, we can see that B could be compensated by the public at large. The factor that blocked liability on the desert island (the effect of paying compensation on A's wellbeing) is no longer present because others can compensate B.

However, even when loss spreading does not determine whether A ought to be compensated, it may still determine who A ought to be compensated by. It may be that if

loss spreading is not possible then compensation is owed by the person who caused the injury (subject to limitations such as that mentioned above), but if it is possible then the wider population ought to bear some of the cost. Or perhaps, especially if the defendant is untraceable or insolvent, the wider population might bear the full cost.³¹ When thinking about the relevance of loss spreading we should keep in mind the distinctness of these two questions: should the victim receive compensation and who should pay? These possibilities also show that the normative questions we are considering, whilst being important to tort law, are broader and more fundamental. They might justify a mixed system of compensation, or even the replacement of tort law with tax-funded compensatory schemes.

Second, imposing a duty to compensate, either on the injurer or the general population, involves intentional harm. It consists in intentionally depriving people of their resources for the sake of others.³² If there is a presumption that intentional harm is worse than unintentional harm, all else being equal, we might think that in cases where harm is caused unintentionally it is better to let losses lie where they fall, rather than impose duties on the injurer or the general population. In addition, forcing people to compensate victims constitutes harming, whereas refraining from compensating victims constitutes failing to benefit. On this basis we may also think that, all else being equal, it is better to let the victim bear the loss than impose duties on the injurer or the public. When it comes to the injurer, at least, these statements are deceptive, as all else is never equal. A number of factors may

³¹ On the distinction between the grounds of recovery and the liability of the injurer, see Jules Coleman, *Risks and Wrongs*, Oxford: Oxford University Press, (1992) at pp. 285 – 7.

³² Jeff McMahan argues that paying compensation is a form of deprivation and therefore involves harm. Imposing a duty to pay compensation, moreover, must count as intentional harm unless the meaning of ‘intentional’ is construed in an implausibly narrow way. McMahan discusses the liability of a tortfeasor, but a similar point can be made about the duties of bystanders to contribute to a victim’s compensation. See Jeff McMahan, “Duty, Obedience, Desert, and Proportionality in War: A Response”, *Ethics*, 122(1), (2011), 135 – 167 at pp. 162 – 3.

count in favour of intentionally imposing a corrective duty on an agent who unintentionally causes harm, thus overriding these presumptions. For example, the harm might be caused wrongfully even if it was unintentional; it might have been avoidable, even if it wasn't wrongful; the injurer might have been the primary beneficiary of the harm-causing act; and the mere fact that harm was *caused* by an agent may matter, even if it was not avoidable. When it comes to the general population, most of these factors are not applicable. The presumptions against intentional harming and causing rather than allowing harm may therefore limit the extent to which we can legitimately force members of the wider community to compensate victims.

Third, the force of loss spreading arguments depends on how widely the loss can be spread. It should be noted that dividing losses is still possible in two party cases: the burden could be divided between the injurer and the victim. This may be an attractive solution if there are no reasons why the injurer should be liable to the victim. Consider a case outlined by Judith Jarvis Thomson in which one person turns on a light switch and due to an unforeseeable malfunction, electrocutes her next door neighbour.³³ Those who believe that merely causing an injury, without culpability or foresight, cannot count in favour of liability will hold that there is no reason why either party should bare the entire loss herself. But if the loss is divided evenly, and the injury is sufficiently severe, this may be too much to expect the injurer to bear. If the loss can be spread across a larger population, however, each person's burden may be small enough to justify its imposition. In this way, the plausibility of loss spreading arguments depends on the size of the population over which the loss is spread.

With these points noted, let us explore the effect of loss spreading in a hypothetical case. Imagine first a two party case taking place on a desert island. The following is a

³³ Judith Jarvis Thomson, *The Realm of Rights*, Cambridge, MASS: Harvard University Press, (1990), p. 229.

standard case of foreseeable but non-culpable harm, in which one party chooses to act for prudential reasons, foreseeably imposing a small risk of harm on another:

Horse Riding: Stephen is riding his horse close to Peter's land. He knows that there is a risk that the horse may damage Peter's fence causing his livestock to escape, so he exercises due care and skill to minimise this risk. Nevertheless, the horse missteps, damages the fence, and Peter's animals escape.

Is there a moral case that Stephen ought to compensate Peter? At least four considerations count in favour of this conclusion. First, the potential harm Peter could suffer is significant. The loss of his livestock will be a serious setback to the wellbeing and resources of himself and his family. Second, Stephen caused harm to Peter. If mere causation matters, this gives Stephen some reason to compensate Peter. Third, Stephen had a better opportunity to avoid imposing the risk than Peter had to avoid being subject to it. Stephen chose to act in a way that imposed a foreseeable risk on Peter. This choice was made for weak prudential reasons: Stephen would not have incurred a large cost in choosing not to impose the risk. Finally, Stephen acts for his sole benefit, and the risk he imposes does not incidentally benefit anyone else.³⁴

If these factors outweigh the countervailing consideration that Stephen is not at fault because he exercised due care and skill, and it is important to allow individuals the freedom to pursue their own ends without the 'moral threat' of liability, this suggests that a standard of strict liability is appropriate in two party cases when the above conditions are satisfied. Luck egalitarians who endorse the distinction between option and brute luck may also be satisfied with this conclusion because Stephen's actions amounted to an informed and calculated gamble, while Peter's loss is a matter of bad brute luck.³⁵ However, the situation

³⁴ See chapter 6 where I consider these factors in greater detail.

³⁵ Ronald Dworkin, *Sovereign Virtue*, Cambridge MA: Harvard University Press, (2000), p. 73.

changes when loss spreading becomes possible. We can imagine the same case occurring in a large society in which the government can tax citizens in order to compensate, either wholly or in part, for Peter's harm. There are at least three arguments one might make to support a fault-based standard of liability when injuries occur in societies large enough to spread losses.

The Anti-Luck Argument: One way anti-luck arguments are often made is by appeal to *ex ante* agreement. In a large society in which the losses involved in accidents can be spread across predefined groups, would it be rational to agree *ex ante* (before one knows whether one will turn out to be a victim, an injurer, both, or neither) to any particular standard of liability for unintentional harm? Jeremy Waldron has put forward a very similar argument to defend a no-fault scheme of compensation,³⁶ but a similar line of reasoning supports the adoption of fault-based over strict liability.

Waldron compares two negligent drivers, Fate and Fortune, both of whom carelessly take their eyes off the road. Luckily for Fortune, no one is in his path while he is distracted. Fate, on the other hand, runs into a pedestrian – an accident that he would have avoided had he kept his eyes on the road. He inflicts serious injuries on the victim and his liability runs into the millions of pounds. Given that Fate and Fortune were equally careless, is it fair, Waldron asks, that Fate should be saddled with massive liability as a result of bad luck while Fortune gets away unencumbered?

Waldron considers two schemes of compensation. The first is liability in negligence, requiring proof that the defendant breached a duty of care causing loss to the claimant. Waldron notes that although this scheme does not alter the drastic inequality of outcome between Fate and Fortune, it has a macabre fairness to it, given that they both incur the same risk of liability. *Ex ante*, Fortune is just as likely to be lumbered with massive liability as Fate, so in acting negligently they are each subject to the same risk of liability. If their

³⁶ Jeremy Waldron, "Moments of Carelessness and Massive Loss", in *Philosophical Foundations of Tort Law*, Ed., David Owen, Oxford: Oxford University Press, (1995).

carelessness creates a 1% risk that they will cause £5 million worth of damage then they each run the same risk – it is just that the risk manifests for Fate but not for Fortune.³⁷ On this view, emphasis on the unequal outcome is misleading. Fate and Fortune are treated fairly because they are both part of the same ‘liability lottery’. The second option canvassed by Waldron is a ‘no fault’ scheme of compensation where charges are levied on all negligent drivers – or all drivers in general – to compensate the victims of accidents. The thought is that on this scheme negligent drivers such as Fate and Fortune are treated fairly because they are each subject to an equal fine, even though few of them will end up causing harm.

Waldron then argues that, given this choice, the *maximin* strategy is most defensible and this supports the latter scheme. Maximin invites us to survey the available options and choose that in which the worst outcome is better than the worst outcome in all the others. This strategy clearly favours the second scheme. The worst outcome, from the driver’s perspective, is that a fine will be levied to fund a scheme of accident compensation. But since all other drivers, or all other negligent drivers, will also pay, the individual amount will be relatively modest. The worst outcome on the first scheme, by contrast, will see those negligent drivers unlucky enough to cause serious injuries burdened with massive liability.

Waldron’s argument captures something important, which applies to *Horse Riding* considered both on the desert island and in the large society, but there are problems with *ex ante* arguments. Arguments that appeal to agreements or decisions made under certain conditions are forms of idealised constructivism. They are constructivist because they hold

³⁷ Of course, it may not be true that they both ran an X% risk of causing harm. Perhaps Fate ran an X+1% risk of causing harm because the traffic was marginally busier that day. This difference is not significant. In acting carelessly on the road, Fate and Fortune had a roughly equal chance of imposing an X% risk or an X+1% risk. The difference in the risks they actually imposed is also a matter of luck. Similar points apply to the degree of damage. Even if both Fate and Fortune caused harm, the degree of harm may vary significantly. Nevertheless, they run a roughly equal risk of causing any particular amount of harm.

that these agreements or decisions fix true judgments about some moral matter, and they are idealised because they appeal to a hypothetical deliberation process subject to moral constraints. Waldron's argument is similar to Rawls' moral contractualism (although Rawls shies away from the claim that the constructivist process generates objectively true conclusions), but other forms of constructivism have been developed.³⁸ As Rawls observes, these thought experiments are hypothetical and nonhistorical. They derive their force from the moral constraints placed on the procedure, which allow them to model moral ideas through the use hypothetical deliberation.³⁹ But these properties generate a worry that the constructivist project is unnecessary. If constructivist arguments depend on the prior validity of the constraints placed on hypothetical deliberation, then surely those constraints can give independent support for moral claims, so why not appeal to these constraints directly?⁴⁰

Constructivist thought experiments, subject to the right moral constraints, may still be valid as decision procedures for reaching correct moral judgements. The second worry, however, is that they also fail in this respect. Francis Kamm has convincingly argued that it

³⁸ See John Rawls, "Kantian Constructivism in Moral Theory", *Journal of Philosophy*, 77, (1980), 512 – 72. For other forms of constructivism, see Thomas Scanlon, *What We Owe to Each Other*, Cambridge, MASS: Harvard University Press, (1998); Christine Korsgaard, *The Sources of Normativity*, Cambridge: Cambridge University Press, (1996) and *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press, (1996); and Roderick Firth, "Ethical Absolutism and the Ideal Observer Theory", *Philosophy and Phenomenological Research*, 12, (1952), 317 – 45.

³⁹ John Rawls, "Justice as Fairness: Political not Metaphysical", *Philosophy and Public Affairs*, 14(3), (1985), 223-251 at pp. 236 – 7.

⁴⁰ Russ Shafer-Landau likens this problem to that presented by Socrates in the *Euthyphro* dialogue. The question there was whether the divine approval of some act constitutes its being pious, or whether divine approval guarantees that an act is pious, but the act's status as pious is not constituted by its being divinely approved. If the latter is the case, then reference to divine approval is unnecessary in analysing the nature of pioussness. See Russ Shafer-Landau, *Moral Realism: A Defence*, Oxford: Clarendon Press, (2003), pp. 42 – 3.

is sometimes impermissible to do something even if this would be in everyone's *ex ante* interest. Kamm cites the following case created by Doug Husak:

All six of us have eaten dinner. We discover that five dishes of soup were lethally poisoned, and one had an antidote in it. We do not know who ate what, and will find out only in an hour when the symptoms appear. We do know that using the organs of the one who ate the antidote is the only way to make a serum to save the other five. Is it permissible to make an agreement now to use one person in this way later?⁴¹

Kamm answers in the negative. The full explanation for this involves a careful summary of her views on moral status and the constraints this status places on the actions of others. But these cases should at least encourage us not to accept *ex ante* arguments uncritically.

Finally, *ex ante* arguments do not capture the significance of causation. This is not problematic if causation has no moral significance, and those who endorse Waldron's argument might accept this. But this is a contentious matter on which it is not safe to make bold assumptions. If causation does matter, it is not clear how its significance should be explained. It may be that moral status and causation are independent threats to *ex ante* arguments.

Luckily, Waldron's insights can be captured without the scaffold of *ex ante* agreements. The motivation for adopting the maximin strategy is to mitigate the effect of bad luck on the fate of negligent actors. The anti-luckist component of Waldron's argument is separable from his contractualism. This might encourage us to consider other ways to mitigate the effect of bad luck on liability. One way to do this is to appeal to the opportunity that agents have to avoid imposing risks. The greater an opportunity a person has to avoid running a risk, the more control she has over her own fate.

⁴¹ Francis Kamm, *Morality, Mortality*, Vol. 2, Oxford: Oxford University Press, (1996), p. 291.

In two party cases such as *Horse Riding* where loss spreading is not possible, there are reasons to opt for a standard of strict liability. The same reasons apply to cases where losses can be distributed amongst broader groups of individuals. It is still true, for example, that the injurer usually has a better opportunity to avoid causing harm than a victim has to avoid suffering it. Nevertheless, there are also reasons why strict liability is undesirable: the risk may be imposed in the course of a permissible and valuable activity, as in *Horse Riding*; the opportunity to avoid the risk may not be very strong; and there will be unfairness, of the kind that exercises Waldron, between those unlucky enough to injure people and those who do not. If loss spreading is an option, those in the *ax ante* position will have the means to mitigate these concerns affording victims of these activities claims against the wider population.

With this safety net in place, fault-based liability may be preferable because we have greater opportunities to avoid being negligent than to injure others without fault. Under a regime of fault-based liability, we can avoid liability by exercising reasonable care whenever we undertake risky activities. If we are unable to exercise such care, we will usually have the option of avoid the activity. Under a regime of strict liability, however, we will be unsure whether engaging in some activity will lead to liability, and although we may be able to choose whether to engage in the activity at all, it will be far too costly to avoid all risk-imposing activities.

There are other reasons why a fault-based standard might be preferable to strict liability in larger societies. In two party cases such as *Horse Riding*, one person's decision to engage in an activity that imposes risks on others is less likely to benefit others compared to risky activities that are performed in larger societies. The permission to drive, build and operate heavy machinery (along with many other activities) plausibly benefits a much wider group of people than those who undertake the activity, or who commission and fund it. This is because these activities sometimes involve large scale social co-operation and contribution

to public goods, such as the building of roads, recreational grounds, the facilitation of transport, and so on.⁴²

These remarks suggest that when broader social benefits are derived from or conditional on permission to perform risky activities, we have less reason to allow those activities to be governed by a regime of strict liability. Moreover, if regimes of strict liability for risky activities in large and interconnected societies proliferate, individuals will accumulate an unacceptable level of liability risk. The prospect of being held strictly liable to compensate others will be accompany any risky activity, and individuals will generally be unable to mitigate this risk, short of avoiding the activity altogether, because liability under a strict regime will be substantially beyond their control.

Let me conclude by stressing that I am not advocating fault-based over strict liability. For all I have argued to the contrary, concerns of distributive justice may favour a general no-fault insurance scheme as advocated by Gregory Keating. The argument is rather that the possibility of loss spreading *at least* shows that fault-based liability is preferable to strict liability (without compulsory insurance). When costs can be borne by the wider community, it is better only to hold people individually liable for their wrongful conduct.

6. Conclusion

This chapter defended the two claims that corrective duties are not a category of fundamental interest and that CJ is not independent of DJ. Corrective duties only apply to harms, or reductions in wellbeing. The factors that determine when a person owes a relational duty to benefit others do not always apply where a victim has been harmed – they are sometimes present when a victim is made non-comparatively badly off. It is the broader

⁴² See *The Benefit Principle* and *The Avoidability Principle* laid out in chapter 5.

category of agent-relative duties that best capture the operation of these factors on a person's duties.

The second claim was defended in a number of ways. The first argument appealed to the reason or duty conferring status of unjust distributive patterns. Given that corrective justice may exacerbate unfair distributions and these distributions give us reason to redistribute, corrective justice theorists cannot shirk the responsibility of justifying the practices of private law in light of these distributive concerns. The next argument, outlined by Gardner, takes a step further. It shows that even if we assume a just distribution, the practice of corrective justice raises further, distinctly distributive concerns.

The final argument takes a step further still. It implies that justifiable standards of liability that are internal to private law are themselves fashioned by distributive concerns. The argument demonstrated this by comparing cases in a 'state of nature', where loss spreading is not possible, with cases where losses can be spread across wider groups. If distributive concerns are irrelevant, we would expect the moral results in both cases to be the same. However, there is reason to believe that this is not the case. The considerations that support strict liability in two party cases are weaker when loss spreading is possible, because adopting a fault-based standard gives greater control to individuals over their own fates and can be combined with a publically funded compensation scheme for victims of agential harms (harm without wrongdoing). Moreover, the benefits of risky activities, as well as the general permission to perform them, tend to be more widely distributed in larger societies.

Conclusion

The arguments contained in this thesis depart from those of most contemporary tort theory in that they are purely normative. One consequence of this approach is that some of the issues we have considered are relatively novel. Two examples of this are the different forms of corrective action considered in chapter 4 and the effect of mistaken belief on actionability considered in chapter 5. My two central claims in those chapters were, first, that rendering a person identically off and equally well off differ in that the latter still involves an interference with the victim, and this has implications for permissibility and the law, and second, that we ought to favour *The Mixed View* over both *The Objective View* and *The Respect View*. The upshot of this conclusion is that we ought to compensate victims when they are harmed but mistakenly believe they are benefitted *and* when they are benefitted but mistakenly believe they are harmed. But although these problems have been little discussed by tort theorists, they are closely connected to broader and more well-known philosophical debates. The difference between harm and interference has been discussed in the context of criminal law,¹ and as I noted in chapter 6, the state's relationship with citizens' mistaken beliefs about their wellbeing is a central issue in political philosophy.

Another consequence of the approach taken in this thesis is that some familiar questions were considered in a different way. Previous theorists of corrective justice have looked to the functioning of private law to understand the nature and importance of corrective duties. As I suggested in chapter 7, private law glosses over the point that duties in response to non-comparatively bad states are not corrective. On a normative approach, the right conception of corrective duties does not depend on what duties are recognised by private law, but on what conception best tracks matters of moral interest. Since duties in

¹ See J. Gardner and S. Shute, "The Wrongness of Rape" in *Oxford Essays in Jurisprudence, Fourth Series*, Ed., J. Horder, Oxford: Oxford University Press, (2000) and Arthur Ripstein, "Beyond the Harm Principle". *Philosophy & Public Affairs*, 34, (2006), 215–245.

response to non-comparatively bad states are also of moral interest, we reached the conclusion that corrective duties are not an importantly distinctive deontic category.

Despite this fundamental difference in approach, I do not mean to suggest that the existing literature has little to contribute to normative questions. Although most philosophers of tort do not consider their theories to be purely normative, they do not consider them purely interpretive either. Principles intended to play a primarily explanatory role – such as wealth maximisation or reciprocity – are worth evaluating on solely moral grounds. This was the inspiration behind chapter 6, where I assessed *The Nonreciprocity Principle*. Although I rejected it, it was an important stepping stone to an alternative set of principles that ground corrective duties. The two central principles advocated were *The Benefit Principle* and *The Avoidability Principle*. Together, these principles explain why individuals often owe corrective duties to those they have harmed.

I do not argue that these principles (as well as considerations of distributive fairness) tell the whole story about the basis of corrective duties, however. Inevitably, this project is left incomplete. Perhaps the biggest omission is a discussion of causation. Any complete theory of tort law must comment on whether causation has any moral significance and if so what theoretical implications this has. Some writers such as Richard Wright, Richard Epstein and Anthony Honoré have hinged their entire theory around causation, while others dismiss it altogether.² Although I have not addressed this explicitly, the resources needed to explain the significance of causation have already been outlined in chapter 6. Let me conclude by briefly explaining why.

² See Richard Wright, "Causation in Tort Law", *California Law Review*, 73, (1985), 1735; Richard Epstein, "A Theory of Strict Liability", *The Journal of Legal Studies*, 2(1), (1973), 151 – 204 and A. M. Honoré, "Responsibility and Luck", *Law Quarterly Review*, 104, (1988), 530. Jeremy Waldron rejects the significance of causation in "Moments of Carelessness and Massive Loss", in David Owen (Ed.), *Philosophical Foundations of Tort Law*, Oxford: Oxford University Press, (1995).

Beforehand, it is important to make two distinctions. The first is between separate concepts that are often bound up with causation such as foreseeability and voluntary agency. Imagine a case where A gets into his car and drives carefully but accidentally runs into B. Some argue that A ought to compensate B, but what is the basis of this argument? We cannot assume, as Epstein does, that causation is the appropriate ground because there is more going on here than mere causation. Although A did not act wrongfully, it was still foreseeable that he would harm B and this distinguishes it from cases of unforeseeable harm. Moreover, A acted voluntarily when he chose to get into his car and this distinguishes it from cases where a person causes harm to another without acting voluntarily. We need to be clear about whether we are saying that mere causation makes a difference, or that it only makes a difference when a person's action stems from their voluntary agency, or when harm is foreseeable.

Second, we need to be clear about the precise implications of causation. Are we saying that causing harm is sufficient to create a compensatory duty? Or are we making the more modest claim that causing harm increases the cost a person must bear in order to compensate her victim? The latter claim leaves open the possibility that the injurer should not bear the *full* cost of compensating the victim, but should share it with the victim or the wider community.

Of course, explaining why causation should be significant at all remains the most important task. To do this, we will need to answer the *anti-luckist* challenge that merely causing harm (independently of foreseeability) is a matter of bad luck, and people should not bear extra burdens as a result of bad luck. Recall now the discussion in chapter 6. *The Benefit Principles* states that whenever A harms B, there is a stronger reason to let the cost fall on the party that benefits more from the risk-creating activity. There is no need, however, to restrict the scope of the principle to the risk-creating activity. It could be extended to cover the harm causing object. If *The Benefit Principle* was expanded in this way, a person's body might fall within its scope. We are all the primary beneficiaries of our own bodies. When our bodies cause harm to others and there are no other relevant facts to determine who should

bear the cost, this fact might count in favour of letting the loss fall on the injurer, as he benefits more from the harm causing object – his body – than his victim.

This sketch is, of course, woefully short. Its purpose was not to make an argument, but to indicate how the ideas explored in previous chapters require a great deal more elaboration and analysis than I have managed here. They may have the potential to further deepen our understanding of this largely undiscovered area of applied moral philosophy: normative tort theory.

Bibliography

Alexander, Larry. "Causation and Corrective Justice: Does Tort Law Make Sense?", *Law and Philosophy*, 6(1), (1987), 1 – 23

Arneson, Richard. "Mill versus Paternalism", *Ethics*, 90, (1980), 470–89

Avraham, Ronen. and Kohler-Hausmann, Issa. "Accident Law for Egalitarians", *Legal Theory*, 12(3), (2006), 181 – 224

Beever, Allan. *Rediscovering the Law of Negligence*, Oxford: Hart Publishing, (2007)

Birks, Peter. "The Concept of a Civil Wrong" in *Philosophical Foundations of Tort Law*, Ed., David Owen, Oxford: Oxford University Press, (1995)

Cane, Peter. *Atiyah's Accidents, Compensation and the Law*, 7th Ed., Cambridge: Cambridge University Press, (2006)

- "Corrective Justice and Correlativity in Private Law", *Oxford Journal of Legal Studies*, 16(3), (1996), 471 – 488

Calabresi, Guido and Melamed, Douglas, "Property Rules, Liability Rules and Inalienability: One View of the Cathedral", *Harvard Law Review*, 85, (1972), 1089 – 1129

- *The Costs of Accidents*, Yale University Press: New Haven and London, (1970)

Chang, Ruth. "The Possibility of Parity", *Ethics*, 112(4), (2002), 659-688

Clayton, Matthew. "The Case Against the Comprehensive Enrolment of Children", *The Journal of Political Philosophy*, 20(3), (2012), 353 – 364

- *Justice and Legitimacy in Upbringing*, Oxford: Oxford University Press, (2006)
- “The Resources of Liberal Equality”, *Imprints*, 5, (2000), 63 – 84

Coase, Ronald. “The Problem of Social Cost”, *Journal of Law and Economics*, 3(1), (1960), 1 – 44

Cohen, G. A. Cohen, *On The Currency of Egalitarian Justice and Other Essays in Political Philosophy*, Princeton: Princeton University Press, (2011)

- “Casting the First Stone: Who Can, and Who Can’t, Condemn the Terrorists”, *Royal Institute of Philosophy Supplements*, 81, (2006), 113 – 136
- “Expensive Taste Rides Again” in *Dworkin and his Critics*, Ed., Justine Burley, Blackwell Publishing, (2004)

Cohen, Joshua. “The Economic Basis of Deliberative Democracy”, *Social Philosophy and Policy*, 6(2), (1989), 25 – 50

Coleman, Jules. *Risks and Wrongs*, Oxford: Oxford University Press, (2003)

- *The Practice of Principle*, Oxford: Oxford University Press, (2001)
- *Markets, Morals and the Law*, Oxford: Oxford University Press, (1998)
- “Tort Law and the Demands of Corrective Justice”, *Indiana Law Journal*, (1992), 349 – 379.
- “The Normative Basis of Economic Analysis: A Critical Review of Richard Posner’s ‘The Economics of Justice’”, *Stanford Law Review*, 34(5), (1982), 1105 – 1131
- “Moral Theories of Torts: Their Scope and Limits: Part 1”, *Law and Philosophy*, 1(3), (1982), 371 – 390
- “Justice and Reciprocity in Tort Theory”, *Western Ontario Law Review*, 14, (1975), 105 – 118

Dancy, Jonathan. *Ethics without Principles*, Oxford: Clarendon Press, (2004)

Dante, Alighieri, *The Inferno, Canto V, Second Circle*, Translated by Henry Wadsworth Longfellow

Duff, Anthony. "Blame, Moral Standing and the Legitimacy of the Criminal Trial", *Ratio*, 23, (2010), 123 – 40

Dworkin, Ronald. *Sovereign Virtue: The Theory and Practice of Equality*, Cambridge, MASS: Harvard University Press, (2000)

- *Law's Empire*, London: Fontana, (1986)
- *A Matter of Principle*, London: Duckworth, (1985)
- "Paternalism," *The Monist*, 56, (1972), 64–84

Epstein, Richard. "Nuisance Law: Corrective Justice and Its Utilitarian Constraints," *The Journal of Legal Studies*, 8(1), (1979), 49 – 102

- Epstein, Richard. "Causation and Corrective Justice: A Reply to Two Critics," *The Journal of Legal Studies*, 8, (1979), 477 – 504
- Epstein, Richard. "A Theory of Strict Liability", *The Journal of Legal Studies*, 2(1), (1973), 151 – 204

Feinberg, Joel. *The Moral Limits of the Criminal Law, Vol. 2, Offence to Others*, Oxford: Oxford University Press, (1988)

- *The Moral Limits of the Criminal Law, Vol. 1, Harm to Others*, Oxford: Oxford University Press (1987)
- *Rights, Justice, and the Bounds of Liberty*, Princeton: Princeton University Press, (1980)

- “Voluntary Euthanasia and the Inalienable Right to Life”, *Philosophy and Public Affairs*, 7, (1978), 93 – 123

Firth, Roderick. “Ethical Absolutism and the Ideal Observer Theory”, *Philosophy and Phenomenological Research*, 12, (1952), 317 – 45

Fletcher, “Fairness and Utility in Tort Theory”, *Harvard Law Review*, 85(3), (1972), 527 – 573

Gardner, John. “What is Tort Law For? Part 2”, (forthcoming)

- Gardner, John. “Torts and Other Wrongs”, *Florida State University Law Review*, (2012), (forthcoming)
- “What is Tort Law For? Part 1. The Place of Corrective Justice”, *Law and Philosophy*, 30, (2011), 1 – 50
- And François Tanguay-Renaud, “Desert and Avoidability in Self-Defense”, 122(1) *Ethics*, (2011), 111 – 134
- “Backwards and Forwards with Tort Law” in *Law and Social Justice*, Eds., J. Keim-Campbell, M. O'Rourke & D. Shier, MIT Press, (2005)
- And S. Shute, “The Wrongness of Rape” in *Oxford Essays in Jurisprudence, Fourth Series*, Ed., J. Horder, Oxford: Oxford University Press (2000)

Goodin, Robert. “Theories of Compensation”, *The Oxford Journal of Legal Studies*, 9(1), (1989), 56 – 75

Hanser, Matthew. “Still more on the Metaphysics of Harm”, *Philosophy and Phenomenological Research* Vol. LXXXII No. 2, (2011), 459 – 469

- “The Metaphysics of Harm”, *Philosophy and Phenomenological Research*, Vol. LXXVII, No. 2, (2008), 421 – 450

Hart, H. L. A. and Honoré, Tony. *Causation in the Law*, 2nd Ed., Oxford: Clarendon Press, (1985)

Hicks, John. "The Foundations of Welfare Economics", *The Economic Journal*, 49, (1939), 696 – 712

Hitchcock, Christopher. "The Role of Contrast in Causal and Explanatory Claims", *Synthese*, 107, (1996), 395 – 419

Hohfeld, Wesley. "Some Fundamental Legal Conceptions as Applied in Judicial Reasoning", *Yale Law Journal*, 23, (1913), 28 – 39

Honoré, A. M.. "Responsibility and Luck: The Moral Basis of Strict Liability", *Law Quarterly Review*, 104, (1988), 530 – 52

Honoré, A. M. "Ownership", in *Oxford Essays in Jurisprudence*, Ed., Anthony Gordon Guest, Oxford: Oxford University Press, (1961)

Hurka, Thomas. *Perfectionism*, Oxford University Press, USA, (1996)

James, "Contractualism's (not so) Slippery Slope", *Legal Theory*, 18(3), (2012), 263 – 292

Kahneman, Daniel. "The Cognitive Psychology of Consequences and Moral Intuition", *The Tanner Lecture in Human Values*, University of Michigan, Ann Arbor, (November 1994).

Kaldor, Nicholas. "Welfare Propositions of Economics and Interpersonal Comparisons of Utility", *The Economic Journal*, 49, (1939), 549 – 552

Kamm, Francis. *Intricate Ethics: Rights, Responsibilities, and Permissible Harms*, Oxford: Oxford University Press, (2007)

- “Rights”, in *The Oxford Handbook of Jurisprudence and Philosophy of Law*, Eds., Jules Coleman and Scott Shapiro, Oxford: Oxford University Press, (2002)
- *Morality, Mortality, Volume 2, Rights, Duties and Status*, Oxford: Oxford University Press, (1996)

Keating, Gregory. “The Priority of Respect over Repair”, *Legal Theory*, 18, (2012), 293 – 337

- “A Social Contract Conception of the Tort Law of Accidents”, in *Philosophy and the Law of Torts*, Ed. Gerald Postema, Cambridge: Cambridge University Press, (2001)
- “Reasonableness and Rationality in Negligence Theory”, *Stanford Law Review*, 48, (1996), 311 – 385

Keren-Paz, Tsaachi. *Torts, Egalitarianism and Distributive Justice*, Ashgate, (2007)

Korsgaard, Christine. *The Sources of Normativity*, Cambridge: Cambridge University Press, (1996)

- *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press, (1996)

Markesinis, Basil., Deakin, Simon and Johnston, Angus. *Markesinis and Deakin’s Tort Law*, 6th Ed., Oxford: Clarendon Press, (2008)

McBride, Nicholas. “Rights and the Basis of Tort Law” in *Rights and Private Law*, Eds., Nolan and Robertson, Oxford: Hart Publishing, (2012)

- McBride, Nicholas and Bagshaw, Roderick. *Tort Law*, 4th Ed., Pearson, (2012)

- McBride, Nicholas. "Thinking about Tort Law – Where Do We Go from Here?", (unpublished ms)

McMahan, Jeff. "Duty, Obedience, Desert, and Proportionality in War: A Response", *Ethics*, 122(1), (2011), 135 – 167

- "The Just Distribution of Harm Between Combatants and Noncombatants" *Philosophy and Public Affairs*, 38(4), (2010), 342 – 379
- "Moral Intuition" in *The Blackwell Guide to Ethical Theory*, Ed., Hugh LaFollette Ed., Blackwell Publishing, (2000)

Nagel, Thomas. *Mortal Questions*, Cambridge: Cambridge University Press, (1979)

Nolan, Donal and Robertson, Andrew. "Rights and Private Law" in *Rights and Private Law*, Eds., Nolan and Robertson, Oxford: Hart Publishing, (2012)

Nozick, Robert. *Anarchy, State, and Utopia*, Wiley-Blackwell: New Ed., (2001)

Parfit, Derek. *On What Matters*, Vol. 1, Oxford: Oxford University Press, (2011)

- "Equality and Priority," *Ratio*, 10, (1997), 202-221
- *Reasons and Persons*, Oxford: Clarendon Press, (1984)

Perry, Stephen. "On the Relationship between Corrective and Distributive Justice", in *Oxford Essays in Jurisprudence*, 237, Ed., Jeremy Horder, 4th series, Oxford: Oxford University Press, (2000)

Posner, Richard. *The Problems of Jurisprudence*, Cambridge, MASS: Harvard University Press, (1990)

- Posner, Richard and Landes, William. *The Economic Structure of Tort Law*, Cambridge, MASS: Harvard University Press, (1987)
- *The Economics of Justice*, Cambridge, MASS: Harvard University Press, (1981)
- “A Theory of Negligence”, *Journal of Legal Studies*, 1, (1972), 29 – 96

Quong, Jonathan. *Liberalism Without Perfection*, Oxford: Oxford University Press, (2010)

Rawls, John *Political Liberalism*, New York: Columbia University Press, (1993)

- “Justice as Fairness: Political not Metaphysical”, *Philosophy and Public Affairs*, 14(3), (1985), 223-251
- “Kantian Constructivism in Moral Theory”, *Journal of Philosophy*, 77, (1980), 512 – 72

Raz, Joseph. “Facing Diversity: The Case of Epistemic Abstinence”, *Philosophy and Public Affairs*, 19(1), (1990), 3 – 46

- *The Morality of Freedom*, Oxford: Clarendon Press, (1988)

Ripstein, Arthur. “As If It Had Never Happened”, *William and Mary Law Review*, 48, (2007), 1957 – 1999

- “Beyond the Harm Principle”, *Philosophy and Public Affairs*, 34, (2006), 215 – 45
- With Benjamin Zipursky, “Corrective Justice in an Age of Mass Torts” in *Philosophy and the Law of Torts*, Ed., Gerald Posteman, Cambridge: Cambridge University Press, (2001)
- *Equality, Responsibility and the Law*, Cambridge: Cambridge University Press, (1999)

Sadursky, Wojciech. “Social Justice and Legal Justice”, *Law and Philosophy*, 3(3), (1984), 329 – 354

Scanlon, Thomas. "Thomson on Self-Defense" in *Fact and Value: Essays on Ethics and Metaphysics for Judith Jarvis Thomson*, Eds. Alex Byrne, Robert Stalnaker and Ralph Wedgwood, Cambridge, MASS: MIT Press, (2001)

- *What We Owe to Each Other*, Harvard: Cambridge, MASS, (2000)

Schaffer, Jonathan. "Contrastive Causation", *Philosophical Review*, 114, (2005), 327 – 358

Shafer-Landau, Russ. *Moral Realism: A Defence*, Oxford: Clarendon Press, (2003)

Shavell, Steven. *Foundations of Economic Analysis of Law*, Cambridge, MASS: Harvard University Press, (2004)

- Shavell, Steven and Kaplow, Louis. *Fairness Versus Welfare*, Cambridge, MASS: Harvard University Press, (2002)

Shiffrin, Seana. "Wrongful Life, Procreative Responsibility, and the Significance of Harm", *Legal Theory*, 5, (1999), 117 – 148

Smith, Stephen. *Contract Theory*, Oxford: Clarendon Press, (2004)

Spencer, John. "Public Nuisance – A Critical Examination", *Cambridge Law Journal*, 48, (1989), 55 – 84

Steel, Sandy. "Private Law and Justice", *The Oxford Journal of Legal Studies*, (forthcoming).

Stevens, Robert. "The Conflict of Rights", in *The Goals of Private Law*, Eds., Robertson and Tang, Oxford: Hart Publishing, (2009)

- Stevens, Robert. *Torts and Rights*, Oxford: Oxford University Press, (2007)

Tadros, Victor. *What Might Have Been*, (forthcoming)

- “Orwell’s Battle with Britain: Vicarious Liability for Unjust Aggression”, (forthcoming)
- “Responses”, *Law and Philosophy* 32 (2-3), (2013), 241-325
- *The Ends of Harm: The Moral Foundations of Criminal Law*, Oxford: Oxford University Press, (2011)
- “Poverty and Criminal Responsibility”, *Journal of Value Inquiry*, 43, (2009), 391 – 413
- *Criminal Responsibility*, Oxford: Oxford University Press, (2005)

Temkin, Larry. *Inequality*, Oxford: Oxford University Press, (1993)

Tennyson, Alfred Lord. *In Memoriam, Part XXVII*, (1850)

Thomson, Judith Jarvis. “More on the Metaphysics of Harm”, *Philosophy and Phenomenological Research* Vol. LXXXII, No. 2, (2011), 436 – 458

- *The Realm of Rights*, Cambridge, MASS: Harvard University Press, (1990)
- *Rights, Restitution, and Risk*, Cambridge: Harvard University Press, (1986)

Vallentyne, “Brute Luck, Option Luck, and Equality of Initial Opportunities”, *Ethics*, 112(3), (2002), 529 – 557

Waldron, Jeremy. “Moments of Carelessness and Massive Loss”, in *Philosophical Foundations of Tort Law*, Ed., David Owen, Oxford: Oxford University Press, (1995)

Weinrib, Ernest. *Corrective Justice*, Oxford: Oxford University Press, (2012)

- *The Idea of Private Law*, Cambridge, MASS: Harvard University Press, (1995)

Williams, Andrew. “How Gifts and Gambles Preserve Justice”, *Economics and Philosophy*, 29(1), (2013), 65 – 85

- “Equality for the Ambitious”, *Philosophical Quarterly*, 53, (2002), 377 – 89

Winfield, Percy. *The Province of the Law of Tort*, Cambridge: Cambridge University Press, (1931)

Wood, Allen. “Humanity as an End in Itself” in Derek Parfit, *On What Matters*, Vol. 2, Oxford: Oxford University Press, (2011)

Wright, Richard. “Causation in Tort Law”, *California Law Review*, 73, (1985), 1735 – 1828

Zimmerman, Michael. “Moral Responsibility and Ignorance”, *Ethics*, 107 (3), (1997), 410 – 426

Zipursky, Benjamin and Goldberg, John. “Torts as Wrongs,” *Texas Law Review*, 88, (2010), 917–86.

- Zipursky, Benjamin. “A Theory of Punitive Damages,” *Texas Law Review*, 84, (2005), 105–72
- “Civil Recourse, Not Corrective Justice,” *Georgetown Law Journal*, 91, (2003), 695–755
- “Philosophy of Private Law” in *The Oxford Handbook of Jurisprudence and Philosophy of Law*, Eds., Jules Coleman and Scott Shapiro, Oxford: Oxford University Press, (2002), 623 – 655
- “Rights, Wrongs, and Recourse in the Law of Torts.” *Vanderbilt Law Review*, 51, (1998), 1–100

Cases

Ashley v Chief Constable of Sussex [2008] UKHL 25

Doherty v Allman (1878) 3 App. Cas. 709, 719 – 20

Fairchild v Glenhaven Funeral Services Ltd [2002] UKHL 22

Hunt v Severs [1994] 2 AC 350 (HL)

Mansfield v Wheetabix Ltd. [1997] EWCA Civ 1352

McFarlane v Tayside Health Board [2000] 2 AC 59

Parry v Cleaver [1970] AC 1 (HL)

R. (on the application of Lumba) v Secretary of State for the Home Department [2011]
UKSC 12

Roberts v Ramsbottom [1980] 1 WLR 823

Sienkiewicz v Grief (UK) Ltd [2011] UKSC 10

Smith v Baker & Sons [1891] AC 325