Introduction

Proof of causation: three types of problem

No proof of causation, no compensation. The general proof rule in perhaps every system of tort law is that the claimant must prove that the relevant aspect of the defendant's conduct was a cause of its injury in order to obtain substantial compensatory damages in respect of that injury.[1] And yet, almost equally universally, legal systems have made exceptions to this rule. This has allowed claimants to succeed either in full or partially, without proof that the defendant's conduct was a cause of their injury.[2]

This book is about the general proof rule and the exceptions that have been made to it.[3] At the outset, it is important to be clear that the book is not principally concerned with the analysis of the concept of causation. Questions about the *nature* of causation arise even when all of the facts are in. The focus of this book is on what the law of tort does and should do in situations where there is evidential *uncertainty* relative to some standard of proof as to whether the phenomenon of interest caused the outcome of interest. This inquiry only gets off the ground once we have decided upon what it is we are seeking to prove, that is, once we have decided upon what it means for something to cause something else.[4]

The aim of the book is to offer analysis and normative critique of the general and exceptional rules, concentrating on English, German, and French tort law, but comparisons will also be made with certain other

---

[1] See Chapter 1, 37–43. The terminology of 'injury', 'loss', and 'damage' is discussed therein: 40–41.
[2] See Chapters 4–6.   [3] The general rules are defined in Chapter 2.
[4] As it is possible, however, to elide (with ensuing confusion) conceptual issues as to the nature of causation with the question of what proof rules should govern the causal issue between the parties, it will be necessary to say something of the boundary between the two later. See Chapter 1, 16–37 and Chapter 5, 239–246.

common law systems.[5] The inquiry is prompted by the many difficult problems to which these rules have given rise. These problems are of three kinds: conceptual problems, problems of consistency, and normative problems.

The main *conceptual* problem in relation to the general rule concerns what is meant by 'proof' of causation. All of the systems examined here refer to an idea of 'probability' in the formulations of their standards of proof. English law requires proof on the balance of probabilities. What, however, is the nature of the probabilistic element in standards of proof? This question arises here because courts are often faced with statistical evidence of causation. The conceptual difficulty is in understanding the extent to which such evidence, and probabilistic evidence more generally, is capable of satisfying legal standards of proof either itself or in combination with other evidence.[6] Some courts have found that statistical evidence on its own is not sufficiently 'individualised' or 'particular' to the case at hand. But the sense in which this is so is often obscure.

Exceptional rules also raise conceptual problems. When legal systems depart from the general proof of causation rule, there are a number of ways of conceptualising this. It may be considered a reversal of the burden of proof. It may be a reduction in the standard of proof. It may be a shift from a threshold, all-or-nothing, system, whereby a person only recovers damages upon proving certain facts to a threshold probability and then recovers in full, to a non-threshold, proportionate system, whereby a person recovers in proportion to the probability of the facts (or some mixture). In the face of uncertainty over causation, some systems even claim to be altering the nature of the concept of causation, or the very basis upon which a person is liable to another in tort law, in order to allow claims to succeed.[7] Properly taxonomising different exceptions is not a mere conceptual obsession. Without a clear understanding of the nature of exceptional rules, we cannot hope to develop the law in a consistent way or to assess its merit.

The main problem of *consistency* is raised by the recognition of exceptional rules alongside the general rule. Here the question is whether any exceptional rule can rationally be contained within narrow limits or whether it undermines insistence upon the general rule *in any case*. For example, is it possible to create a localised system of proportionate

---

[5] The other comparisons are principally with Canada (see Chapter 4) and the United States (see Chapters 4–6).
[6] See Chapter 2.   [7] See Chapter 4, 166–167; Chapter 5, 225ff.

liability without undermining the threshold, all-or-nothing, general regime? The history of the development of exceptional rules in each system is marked by examples of an exceptional rule being created, its potential radically to undermine the general rule then being recognised, before, finally, it is subject to some apparently arbitrary limitation.[8] Some have argued that there is a kind of necessity to this history, and that only arbitrary limitations can *ever* be found once the initial departure from the general rule is made.[9] This book argues, on the contrary, that principled exceptions can be made, even if all systems have imposed arbitrary restrictions upon their current exceptional rules.[10] At any rate, it is a pressing question for all systems whether (and how) their exceptional rules can be rationalised consistently with their commitment to applying the general rule.

The general and exceptional rules give rise to two connected *normative* problems. First, there is the question of whether the general rule can be justified. The universal recognition of (some form of) exceptional rule(s) challenges the normative foundations of the general rule. Perhaps it is misguided. Perhaps the burden of proof on causation ought quite generally to be upon defendants. Perhaps uncertainty over causation should always be reflected simply by damages proportionate to the probability of causation, rather than by a threshold, all-or-nothing rule. Second, there is the question of whether, if the general rule can be justified, its justifications either positively suggest or leave open the possibility of exceptional rules being created (and, if so, what the scope of such rules should be). If one thought, for example, that the general rule were justified as efficient only in situations in which defendants cannot reliably predict that claimants could not prove causation against them, this might provide (an incentives-based) reason to create exceptional rules in situations where proof of causation is predictably and recurrently impossible.[11]

The book's argument

This book argues that the common law's version of the general rule – the requirement that the claimant prove that the relevant aspect of the defendant's conduct was a cause of its injury on the balance of

---

[8] See, especially, Chapter 5, 237–239.

[9] J. Morgan, 'Causation, Politics and Law: the English – and Scottish – Asbestos Saga', 63.

[10] See Chapter 4, 174–197.   [11] See Chapter 3, 135–137.

probabilities in order to obtain compensatory damages in respect of that injury – is generally justified.[12]

It is also argued, however, that a set of exceptions to this rule can be justified. The argument is that these exceptions should primarily be governed by the following principles:

> *Reliance upon wrongful conduct principle*. A defendant is not permitted to rely upon another's wrongful conduct in order to avoid a liability to pay damages, which would otherwise arise against that defendant, where that has the effect of depriving a person, whose injury would not have occurred had no-one behaved wrongfully in relation to that person, of any right to compensation in respect of that injury.[13]
>
> *Prevented claim principle*. If D1 has either wrongfully injured C or wrongfully prevented C from recovering damages for a wrongful injury caused by D2, C should recover damages from D1, even if C is unable to establish whether D1 wrongfully caused the injury or prevented C's recovery from D2.[14]
>
> *Proven causation principle*. If the evidence shows that D has wrongfully caused x per cent of the injuries within a group of persons, each of whom has suffered an injury which may have been caused by D, but none of whom can establish on the balance of probability that D's wrongful conduct was a cause of their injury, D should be liable either (a) to each member for x per cent of each member's losses or (b) to contribute in the entirety of the amount of loss it has caused to a court held fund from which members may claim proportionate shares, to the extent that this does not more likely than not require D to pay for more loss than D has wrongfully caused.[15]

The fundamental justification of the first principle is tort law's concern to ensure that a person is not deprived of a right to compensation in respect of an injury which would not have occurred had no one behaved wrongfully in relation to that person. The fundamental justification of the other principles is that requiring the defendant to pay compensation is in fact the best way to enforce its secondary moral duty of corrective justice in the special circumstances to which these principles apply. For example, suppose that D1 and D2 have each caused an injury to C1 and C2, but it cannot be determined whether C1's injury was caused by D1 or D2 nor whether C2's injury was caused by D1 or D2. Here, it is more consistent with each defendant's secondary moral duty

---

[12] See Chapter 3.   [13] See Chapter 4, 189–192.

[14] See Chapter 4, 175–185. The *reliance upon wrongful conduct principle* is partly justified by the *prevented claim principle*, but it is broader than the latter principle.

[15] See Chapter 6, 357–359.

to pay compensation to require them to pay damages into a fund to be dispersed to their victims than to deny liability. These principles justify liability, it will be argued, in what will be termed the indeterminate defendant situation, the claimant indeterminacy situation, and complex variants of these situations.[16] In short, the book only countenances exceptional rules being created in circumstances where either the defendant has in fact wrongfully caused injury or the claimant has been the victim of an injury that would not have occurred without wrongful conduct.[17]

Beyond these principles and these situations, the book argues, legal systems should not go. It thus rejects, for example, a reversal of the burden of proof on causation where the defendant has behaved with gross fault, a general reversal of the burden of proof, and proposals to abandon an all-or-nothing system of proof in favour of a system of proportional, probabilistic liability in relation to causation. It also rejects the award of substantial compensatory damages purely in respect of a lost chance, which is sometimes (incorrectly) termed a proportional liability.[18]

Even if the specific arguments of the book are not accepted, it also aims, more modestly, to present the conceptual and normative issues at stake more clearly than before, partly by drawing upon tort theory and the theory of the law of evidence, and partly through critically sifting the normative repository which the case law and legal literature of different legal systems provide. In this way, it will at least show some responses to causal uncertainty in tort law to be dead ends.

Causal uncertainty: its nature and sources

The kinds of causal uncertainty faced by tort law can be classified in different ways for different purposes. It will be helpful straight away to introduce two particular classifications of causal uncertainty which are commonly found in legal doctrine and legal literature and will be discussed later in the book.[19] These are the distinction between uncertainty over *general* and *specific* causation and the distinction between *factual* and *scientific* uncertainty. Both of these, broadly, draw distinctions

---

[16] See Chapters 4, 5, and 6.
[17] These are stated here as necessary, not sufficient, conditions of an exceptional rule being created. These necessary conditions follow from applying only the three principles stated.
[18] See Chapter 6.   [19] See Chapter 2, 67ff.

as to the *nature* of the uncertainty faced by the law. A third taxonomy, which distinguishes causal uncertainty by reference to the *source* of the uncertainty, is then introduced.

Uncertainty over general and specific causation

Causal uncertainty in tort law is often said to relate either to *general* causation or to *specific* causation.[20] The former concerns uncertainty over whether c can cause e (Can smoking cause cancer?). The latter concerns uncertainty over whether c caused e on a specific occasion (Did smoking cause John's cancer?). In essence, there is a broad distinction between cases where our uncertainty extends to ignorance over whether c can ever cause e and cases where our uncertainty is simply over whether c caused e on a specific occasion.

Some nuances of the general–specific distinction are explored later.[21] Two points should, however, be borne in mind now. First, one should avoid the impression that general causation is of any interest to tort law except in so far as it affects the fact-finder's judgement on specific causation. Tort law's primary concern is whether the alleged cause actually caused the outcome in the case before the court, not whether it can generally do so. Second, it must be emphasised that uncertainty over general causation is not wholly distinct from uncertainty over specific causation. In so far as we are uncertain about the former, we may be uncertain about the latter. And, conversely, in so far as a claimant has provided evidence that c can cause e in circumstances similar to its own, the claimant has already adduced *some* evidence that c did cause e.[22]

Factual and scientific causal uncertainty

Courts and legal academics often distinguish – with various purposes in mind – between 'factual' and 'scientific' uncertainty over causation.[23] Without more, this is an unsatisfactory distinction. Scientific uncertainty is just uncertainty over facts. However, the distinction can be used in ways which make it possible to attach some sense to it.

First, it may be used to isolate a set of typical problems that may arise when courts are faced with technical evidence of causation, presented by experts – evidence which will relate to areas of knowledge going beyond

---

[20] See, for example, L. Khoury, *Uncertain Causation in Medical Liability*, 49–50.
[21] See Chapter 2, 67ff.   [22] See Chapter 2, 94ff.
[23] See, for example, L. Röckrath, *Kausalität, Wahrscheinlichkeit und Haftung*, 213; L. Khoury, *Uncertain Causation*, 47–8.

common general knowledge.[24] These include uncertainty as to the probative force of *statistical* evidence, uncertainty arising from *disagreement* between scientific experts over the existence of a causal relationship, and uncertainty over the *reliability* of scientific evidence.[25]

Second, the distinction may also be used simply to refer to the difference between uncertainty over *general* causation and *specific* causation. Sciences, it might be said, are in the business of discovering causal generalisations or laws; they are not in the business of discovering whether a particular element of one of those generalisations was instantiated at a particular time.[26] This might be generally true, though some sciences are concerned with whether a particular fact was instantiated on a particular occasion, for example, forensic sciences such as ballistics.

Even in so far as a distinction between factual and scientific uncertainty can be said to refer to some conceptually robust, or even pragmatically robust, distinction, any attempt to rely upon it for some normative purpose, as some courts have done, must be thoroughly scrutinised.[27]

Causal uncertainty by source

It is possible to classify causal uncertainty by reference to typical causes of it. To say that c is a cause of causal uncertainty is, roughly, to say that without c, we would know or have greater understanding of the causal facts and normally we would know the causal facts were it not for c.

Time

The passage of time is a cause of causal uncertainty. Memories fade, records are lost, and people die. Suppose C used a drug forty years ago which only now materialises in injury, and that drug was manufactured by different manufacturers at the time it was ingested.[28] Uncertainty may arise over the identity of the manufacturer because C or C's pharmacist or doctor no longer has a record or recollection of the manufacturer. Had less time elapsed, this may well have been otherwise.

[24] See, using it in this sense, L. Khoury, *Uncertain Causation*, 47.
[25] This is not, of course, to claim that these types of problem require reasoning which is not continuous with general inductive reasoning. See D. Dwyer, *The Judicial Assessment of Expert Evidence*.
[26] See, emphasising the role of science in coming to know causal generalisations, R. Wright, 'The NESS Account of Natural Causation: A Response to Criticisms', 289.
[27] See Chapter 5, 237–238; 282–285.
[28] As in *Sindell* v. *Abbott Laboratories* (1980) 26 Cal 3d 588. See Chapter 4, 165–168.

Multiplicity of possible causes

If there is only one possible cause of an effect, then it is obviously easier to determine whether the defendant's conduct brought about the effect than if there are multiple possible causes of that effect. So the multiplicity of causes of an effect may lead to causal uncertainty.

Similarity of possible causes

If there are different circumstances (different causally sufficient sets of conditions) which can bring about e, and those circumstances operate in a similar way to bring about e, this often causes uncertainty as to which set of circumstances actually caused e. The classic example of this situation is where two or more hunters negligently fire in the direction of the claimant, at the same time, using identical bullets, and the claimant suffers injury from one bullet. The similarity of the causal process by which any of the hunters could have brought about the claimant's injury is one significant reason why we are uncertain about which hunter did so.[29]

Unobservability of causation

Many causal processes are directly observable.[30] Burnings, breakings, smashings, liftings are all such. But many causal processes (mechanisms) produce their effects behind closed doors. Most obviously, many mechanisms which bring about disease occur within the human body. One reason why it is difficult to attribute the causation of a mesothelioma (an asbestos-induced cancer) to a particular exposure to asbestos is the unobservability of the mechanism by which asbestos produces mesothelioma.

Absence of mechanistic knowledge

More generally, our absence of knowledge about a causal mechanism can create uncertainty. Sometimes we may be very confident that c can cause e, yet understand little about the mechanism, as is the case with asbestos and mesothelioma.[31] If we want to discern whether c rather than any other number of potential causes of e actually caused e, we often need information about the mechanism by which c does this. For example, if we know that a mechanism operates by leaving a certain residue on a person's lung, then this may be crucial evidence in earmarking c as a

---

[29] See Chapter 4, 161–165; 170–174.   [30] See Chapter 2, 67.   [31] See Chapter 5, 266ff.

cause. The consequence of not having such knowledge is that it is often only possible to speak in terms of statistical probabilities as to whether c was a cause of e.[32]

Counterfactual nature of causation

Assessing whether c was a cause of e often, possibly always, involves answering a counterfactual question: had c not occurred would e have occurred?[33] It is often said that the hypothetical nature of such a question creates uncertainty.[34] This is not as obvious as it may seem. Many hypotheticals (counterfactuals) are evidentially straightforward: 'would I have won Wimbledon if I had entered last year?' or 'would I have donated money to Oxfam if I had known they used it all to buy champagne?'

However, if we accept that it is generally easier to gain evidence as to observable events, we should also accept that it is generally more difficult to assess counterfactuals. It is not possible to observe whatever it is that makes a counterfactual true. We do not *observe* what would have happened, only what did happen.

Indeterminism and randomness

Indeterminism leads to causal uncertainty.[35] A world is indeterministic in so far as its state at one time does not entail (by virtue of physical laws) its state at another time.[36] If we imagine a world which is exactly the same in all respects at t, it is possible that the world will develop in different ways after t if that world is indeterministic. If our world is indeterministic, then there is an ineliminable uncertainty built into its fabric, because if a process is indeterministic, we could never *in principle* be sure how it would have developed (knowing all the facts will not help). In so far as we are concerned with an indeterministic process, we must talk in terms of probability.

In reality, we have little idea which, if any, aspects of our world are deterministic or indeterministic.[37] Many events seem to be undetermined (such as which lottery ball will appear next or human decisions). But there is always the possibility that some hidden difference between initial conditions explains a difference in outcomes. Nonetheless, we

---

[32] See Chapter 2, 62.   [33] See Chapter 1, 16–17.
[34] See, for example, R. Strassfeld, 'If. . . : Counterfactuals in the Law', 348–52 (discussing 'counterfactual dread').
[35] See, G. Mäsch, *Chance und Schaden*, 19–20.
[36] See C. Hoefer, 'Causal Determinism', section I.   [37] See Chapter 6, 351, n 224.

may sensibly say that many events display *randomness*, where random-ness consists in variation of outcomes despite extremely similar initial conditions. This randomness means that we do not have specific evidence how a certain process would have developed on a specific occasion; again, the only evidence is probabilistic.

Human agency

Human beings can create causal uncertainty. Suppose that C is negligently shot by either D1 or D2. D1 and D2 use different kinds of bullets and the bullet which missed C cannot be found. A surgeon, told that the bullet lodged in C's body is essential evidence, carelessly misplaces it. Had the surgeon not behaved in that way, we would know which of D1 or D2 negligently injured C.

Cost

Sometimes we do not have causal knowledge because it is very expensive to obtain. For example, in the Canadian case of *Letnik* v. *Municipality of Metropolitan Toronto*, the claimant's ship suffered minor damage as a result of the defendant's negligence.[38] The ship unexpectedly sank two weeks later. The cost of raising the ship would have been somewhere between \$324,000 and \$890,000 (while the ship was itself valued at around \$450,000). If the ship had been raised, it would in all likelihood have been possible to ascertain whether the initial collision, resulting from the defendant's negligence, had caused the ship to sink.

Tort law

In this book, 'tort law' refers primarily to the bodies of law which provide persons with remedies which respond to breaches of legal duties owed to them by other persons, where those duties are owed independently of contracts between them or declarations of trusts. Analytically, if X breaches a duty owed to Y, X also violates Y's legal rights. Conversely, whenever Y's legal rights are violated by X, X must have breached a legal duty owed to Y.[39] One could equally, then, describe the subject matter of this book as the bodies of law which provide remedies to persons that

---

[38] (1988) 44 CLT 69.

[39] At least on a plausible Hohfeldian analysis of legal rights and duties: see W. Hohfeld, 'Some Fundamental Legal Conceptions as Applied in Judicial Reasoning'. For similar views, R. Stevens, *Torts and Rights*, 2; E. Descheemaeker, *The Division of Wrongs*, 22.