

Chicago-Kent College of Law

Scholarly Commons @ IIT Chicago-Kent College of Law

All Faculty Scholarship

Faculty Scholarship

December 1985

Causation in Tort Law

Richard W. Wright

IIT Chicago-Kent College of Law, rwright@kentlaw.iit.edu

Follow this and additional works at: https://scholarship.kentlaw.iit.edu/fac_schol



Part of the [Torts Commons](#)

Recommended Citation

Richard W. Wright, *Causation in Tort Law*, 73 Cal. L. Rev. 1735 (1985).

Available at: https://scholarship.kentlaw.iit.edu/fac_schol/697

This Article is brought to you for free and open access by the Faculty Scholarship at Scholarly Commons @ IIT Chicago-Kent College of Law. It has been accepted for inclusion in All Faculty Scholarship by an authorized administrator of Scholarly Commons @ IIT Chicago-Kent College of Law. For more information, please contact jwenger@kentlaw.iit.edu, ebarney@kentlaw.iit.edu.

Causation in Tort Law

Richard W. Wright

TABLE OF CONTENTS

I.	THE SCOPE AND SIGNIFICANCE OF THE CAUSAL INQUIRY IN TORT LAW	1741
A.	<i>The Scope of the Causal Inquiry: Distinguishing the Causal Inquiry from the Tortious-Conduct and Proximate-Cause Inquiries</i>	1741
1.	<i>Malone</i>	1742
2.	<i>Hart and Honoré</i>	1745
3.	<i>Epstein</i>	1750
B.	<i>The Significance of the Causal Inquiry: Linking the Causal Inquiry and the Tortious-Conduct Inquiry Through the Tortious-Aspect Causation Requirement</i>	1759
1.	<i>Green's Duty-Risk Theory</i>	1761
2.	<i>Keeton's "Harm Within the Risk" Theory</i>	1763
3.	<i>The Tortious-Aspect Causation Requirement</i>	1766
4.	<i>The Power of the Tortious-Aspect Causation Requirement</i>	1771
II.	THE NATURE AND CONTENT OF THE CAUSAL INQUIRY ...	1774
A.	<i>The But-For Test and Its Limits: Duplicative and Preemptive Causation</i>	1775
B.	<i>Efforts to Modify the But-For Test</i>	1777
1.	<i>Detailing the Manner of Occurrence</i>	1777
2.	<i>Detailing the Injury</i>	1778
3.	<i>Excluding Hypothetical Facts</i>	1780
4.	<i>Aggregating Multiple Potential Causes</i>	1780
C.	<i>The Substantial-Factor Formula</i>	1781
D.	<i>Undefined, Directly Observable Causal Contribution</i>	1784
E.	<i>The NESS (Necessary Element of a Sufficient Set) Test</i> .	1788
1.	<i>The Philosophic Basis of the NESS Test</i>	1789
2.	<i>The Duplicative-Causation Cases</i>	1791
3.	<i>The Preemptive-Causation Cases</i>	1794
4.	<i>Distinguishing the Damages Issue: The Successive- Injury and Overwhelming-Force Cases</i>	1798
5.	<i>Theft, Nonuse, or Misuse of Defective or Missing Safety Devices</i>	1801
F.	<i>The Factual Nature of the Causal Inquiry</i>	1803

1735

1. <i>The Hypothetical-Inquiry Argument</i>	1803
2. <i>Malone's Arguments</i>	1807
III. CAUSATION, PROBABILITY, AND RISK: PROBLEMS ON THE FRONTIERS OF TORT LIABILITY	1813
A. <i>The Reduced-Chance and Increased-Risk Cases</i>	1814
B. <i>The Alternative-Causation Cases</i>	1816
C. <i>Distinguishing Probability Statements from Causal Attribution</i>	1821
CONCLUSION	1826

Causation in Tort Law

Richard W. Wright†

In all of tort law, there is no concept which has been as pervasive and yet elusive as the causation requirement, which relieves a defendant of liability if his tortious conduct was not a cause of the plaintiff's injury.¹ Although described by no less an authority than William Prosser as one of the "simplest and most obvious" problems in determining tort liability,² the causation requirement has resisted all efforts to reduce it to a useful, comprehensive formula and has been the subject of widely divergent views concerning its nature, content, scope, and significance.

Through the early part of this century, judges applied causal analysis very broadly. They used causal rhetoric to determine not only whether the defendant's tortious conduct had contributed to an injury but also whether the defendant should be held legally responsible for the injury in the light of the other contributing factors. This broad use of causal language was sharply challenged during the second quarter of this century by the legal realists and their colleagues. They insisted that the only causal issue in a tort case is the simple factual issue of actual causation or causation-in-fact: whether the defendant's conduct actually contributed to the plaintiff's injury. They noted that most of the courts' allegedly causal analysis had nothing to do with the issue of causation-in-fact, but rather, under the misleading label of "proximate cause" or "legal cause," served as a smokescreen for policy judgments on whether the defendant ought to be held liable for an injury to which his conduct had clearly contributed. The actual causation question was considered to be of little help on this issue of ultimate responsibility, since there are innumerable causes of each injury. Instead, the realists argued that, in

† Associate Professor of Law, Chicago-Kent College of Law, Illinois Institute of Technology. B.S. 1968, California Institute of Technology; J.D. 1973, Loyola University, Los Angeles; LL.M. 1976, Harvard University. I would like to thank George Christie, Donald Marshall, Paul Shupack, Stewart Sterk, Richard Weisberg, and Elliott Weiss for their helpful comments on the initial draft of this article.

1. E.g., O.W. HOLMES, THE COMMON LAW 64 (M. Howe ed. 1963); W. PROSSER, HANDBOOK OF THE LAW OF TORTS § 41, at 236-37, 241 (4th ed. 1971). This statement of the causation requirement is not meant to exclude injunctive relief against threatened injury. See *id.* § 1, at 2. The statement also is ambiguous with respect to whether the causation requirement applies to the defendant's conduct as a whole or to the tortious aspect of the conduct. This ambiguity is addressed *infra* in Part I, Section B.

2. W. PROSSER, *supra* note 1, § 41, at 237; accord L. GREEN, RATIONALE OF PROXIMATE CAUSE 4, 132-33, 135-36 (1927); 2 F. HARPER & F. JAMES, THE LAW OF TORTS § 20.2, at 1110-21 & n.4 (1956 & Supp. 1968). Harper and James's chapter 20 is a reprint of James & Perry, *Legal Cause*, 60 YALE L.J. 761 (1951).

order to reach the socially advantageous, expedient, or just result in each case, the policy issues should be openly addressed under the more appropriate headings of "duty," "interests protected," "damages," and so forth.³

Although the realists were unable to eliminate the use of proximate-cause language, they were very successful in gaining academic and judicial acceptance of their view that causation-in-fact is the only truly causal issue, and that proximate-cause determinations are noncausal policy judgments on the appropriate limits of liability for actually caused harm.⁴

However, courts and legal scholars were unable to formulate a comprehensive, policy-neutral account of actual causation.⁵ These failures eventually led to the view that the actual causation issue itself is permeated by policy considerations. This view was expressed most forcefully in an influential article published in 1956 by Wex Malone. He asserted that the actual causation requirement is just one more doctrinal concept that is manipulated by judges to further the relevant socially preferred policy in each case.⁶ His view has been widely accepted,⁷ even in limited form for the most difficult cases by some of the staunchest defenders of actual

3. See L. GREEN, JUDGE AND JURY 29-37, 191, 195-96, 222-25, 230-31, 242-43 (1930); L. GREEN, *supra* note 2; W. PROSSER, *supra* note 1, §§ 41, 42, at 236-37, 239, 244-50 (4th ed. 1971); *id.* §§ 45, 46, at 311-21 (1st ed. 1941); see F. HARPER, A TREATISE ON THE LAW OF TORTS § 110, at 257-58 (1933); Edgerton, *Legal Cause* (pt. 2), 72 U. PA. L. REV. 343, 373 (1924); Gregory, *Proximate Cause in Negligence—A Retreat from "Rationalization,"* 6 U. CHI. L. REV. 36 (1938); Morris, *On the Teaching of Legal Cause*, 39 COLUM. L. REV. 1087 (1939).

4. See, e.g., G. CHRISTIE, CASES AND MATERIALS ON THE LAW OF TORTS 245-46, 264 (1983); R. EPSTEIN, C. GREGORY & H. KALVEN, CASES AND MATERIALS ON TORTS 272-73, 309 (4th ed. 1984) [hereinafter cited as EPSTEIN, TORTS]; J. FLEMING, THE LAW OF TORTS 170-71, 179-80 (6th ed. 1983); M. FRANKLIN & R. RABIN, CASES AND MATERIALS ON TORT LAW AND ALTERNATIVES 269, 302 (3d ed. 1983); 2 F. HARPER & F. JAMES, *supra* note 2, §§ 20.1, 20.2, 20.4; C. MORRIS & C.R. MORRIS, MORRIS ON TORTS 154-201 (2d ed. 1980); W. PROSSER, *supra* note 1, §§ 41, 42, at 236-37, 239, 244-50.

5. See *infra* text accompanying notes 166-225.

6. Malone, *Ruminations on Cause-in-Fact*, 9 STAN. L. REV. 60 (1956).

7. See, e.g., Cole, *Windfall and Probability: A Study of "Cause" in Negligence Law* (pts. 1 & 2), 52 CALIF. L. REV. 459, 764 (1964); Delgado, *Beyond Sindell: Relaxation of Cause-in-Fact Rules for Indeterminate Plaintiffs*, 70 CALIF. L. REV. 881, 891-92 (1982); Pedrick, *Causation, the "Who Done It" Issue, and Arno Becht*, 1978 WASH. U.L.Q. 645; Pound, *Causation*, 67 YALE L.J. 1 (1957); Robinson, *Multiple Causation in Tort Law: Reflections on the DES Cases*, 68 VA. L. REV. 713, 713-14 (1982); Rosenberg, *The Causal Connection in Mass Exposure Cases: A "Public Law" Vision of the Tort System*, 97 HARV. L. REV. 849, 855 n.27, 863-64 (1984); Strachan, *The Scope and Application of the "But For" Causal Test*, 33 MOD. L. REV. 386, 389-95 (1970); Weinrib, *A Step Forward in Factual Causation*, 38 MOD. L. REV. 518, 529-33 (1975); Zweir, *"Cause in Fact" in Tort Law—A Philosophical and Historical Examination*, 31 DE PAUL L. REV. 769 (1982); see 2 F. HARPER & F. JAMES, *supra* note 2, § 20.2 & nn.16 & 17 comments at 92-93, 94-95 (Supp. 1968); W. KEETON, D. DOBBS, R. KEETON & D. OWEN, PROSSER AND KEETON ON THE LAW OF TORTS §§ 41, 42, at 264-65 & n.7, 279-80 (5th ed. 1984) [hereinafter cited as PROSSER & KEETON]; cf. C. MORRIS & C.R. MORRIS, *supra* note 4, at 188-90 (asserting that liability is sometimes imposed for policy reasons although causation is lacking).

causation as a factual inquiry.⁸ It has been adopted enthusiastically by the economic analysts of tort law, who argue that the socially preferred policy in every case is the maximization of social wealth.⁹

But there has also been another, very different reaction to the views of the legal realists, which was initiated by H.L.A. Hart and Tony Honoré's landmark book, *Causation in the Law*, originally published in 1959.¹⁰ As an alternative to the ad hoc social policy analysis of the legal realists, Hart and Honoré attempted to rehabilitate the broad use of causal analysis to address questions of ultimate liability. They argued that the ordinary, nonlegal use of causal language provides definite, commonsense causal principles that can be, should be, and are used by the courts to frame factual inquiries which encompass not only the narrow issue of cause-in-fact but also most of the proximate-cause issues.¹¹ Indeed, they stated, "causing harm constitutes not only the most usual but the primary type of ground" for imposing liability.¹² Subsequent tort scholars, most notably Richard Epstein, have attempted to elaborate on this statement.¹³

8. See A. BECHT & F. MILLER, THE TEST OF FACTUAL CAUSATION IN NEGLIGENCE AND STRICT LIABILITY CASES 24, 44 n.57, 78 n.109, 81-82 & n.114, 85-87, 99, 104-06 & n.157, 112 n.173, 118, 120, 124-26, 128-30, 135-37, 221 (1961); J. FLEMING, *supra* note 4, at 172-73; H.L.A. HART & T. HONORÉ, CAUSATION IN THE LAW lxi-lxii, 7-8, 62, 101-02 & n.46, 239-41, 410, 412-16 (2d ed. 1985); Green, *The Causal Relation Issue in Negligence Law*, 60 MICH. L. REV. 543, 548-49, 553-55 & n.29, 560-61, 568 (1962); see also W. PROSSER, *supra* note 1, § 41, at 243 & n.53.

9. See, e.g., Calabresi, *Concerning Cause and the Law of Torts*, 43 U. CHI. L. REV. 69, 69 n.1, 86-87, 105-08 (1975); Grady, *A New Positive Economic Theory of Negligence*, 92 YALE L.J. 799, 799-800, 804 n.21 (1983); Landes & Posner, *Causation in Tort Law: An Economic Approach*, 12 J. LEGAL STUD. 109, 110-11, 134 (1983); Shavell, *An Analysis of Causation and the Scope of Liability in the Law of Torts*, 9 J. LEGAL STUD. 463, 464, 502-03 & n.79 (1980). For my critique of the economic analysts' views on causation in tort law, see Wright, *Actual Causation vs. Probabilistic Linkage: The Bane of Economic Analysis*, 14 J. LEGAL STUD. 435 (1985).

10. H.L.A. HART & A. HONORÉ, CAUSATION IN THE LAW (1st ed. 1959). The recently published second edition of this book contains a lengthy new preface but otherwise is generally the same as the first edition. H.L.A. HART & T. HONORÉ, *supra* note 8. Two changes relevant to the issues addressed in this Article are discussed *infra* at notes 12 & 31.

11. H.L.A. HART & A. HONORÉ, *supra* note 10, at 3-5, 23-25, 61-63, 65, 83-105, 123-25, 261-62, 273-76; H.L.A. HART & T. HONORÉ, *supra* note 8, at xxxiv-xxxvi, lii-liii, 3-5, 24-27, 65-67, 69-70, 88-111, 130-32, 291-92, 304-07. This objective is stated so emphatically and repeatedly throughout the book that it is difficult to comprehend Izhak Englad's failure to see it and his criticism of John Borgo, who did see it, for "misunderstanding" Hart and Honoré. Englad, *The System Builders: A Critical Appraisal of Modern American Tort Theory*, 9 J. LEGAL STUD. 27, 57 n.135 (1980) (criticizing Borgo, *Causal Paradigms in Tort Law*, 8 J. LEGAL STUD. 419, 421-25 & n.17 (1979)). Apparently, Englad does not distinguish between the prevailing academic policy-oriented interpretation of proximate cause, which Hart, Honoré, and Borgo were attacking, and the commonsense judgments independent of particular legal policies which Hart, Honoré, and Borgo argue are controlling in actual adjudication.

12. H.L.A. HART & T. HONORÉ, *supra* note 8, at 65; see *id.* at lxxvii-lxxxi, 66-67, 302. However, in other passages (especially in the second edition of their book), Hart and Honoré state that causing harm usually must be coupled with wrongful conduct to create liability. *Id.* at xxxv, xlivi-xlvii, lxxv-lxxvii, 131-32, 254-55, 466.

13. A. HARARI, THE PLACE OF NEGLIGENCE IN THE LAW OF TORTS 49-56 (1962); Epstein, *A*

In this Article, I present a systematic critique of these various accounts of the causation requirement in tort law,¹⁴ and, building on that critique, I construct a comprehensive new account which hopefully will be more satisfactory than any of the earlier accounts.

This Article is divided into three parts. Part I discusses the manner in which the causal inquiry fits into tort liability analysis. Part II discusses the nature and content of the causal inquiry itself. Part III discusses some current problems and controversies dealing with the relationships among causation, probability, and risk.¹⁵

More specifically, I argue in Part I that much of the disagreement and confusion in the debates about causation stems from failures to distinguish the causal inquiry from other elements of the courts' liability analysis, or to pay sufficient attention to the links that connect the different elements. As a result of these failures, most writers substantially overrate or underrate the role that the causal inquiry plays in tort liability analysis. Thus, some writers attribute to the causal inquiry issues that are properly associated with the tortious-conduct or proximate-cause inquiries. Conversely, other writers trivialize the causal inquiry by linking it to the actor's conduct as a whole, rather than focusing it—as the courts do—on the tortious aspect of the actor's conduct, and seek to fill the resulting void with inadequate duty or "harm within the risk" analyses. When the causal inquiry is properly limited and focused, it is seen to play a powerful but not conclusive role in determining tort liability.

Part II begins with a brief discussion of the traditional "but for" (necessary condition) test of causation and its inability to account for the overdetermined-causation cases—cases in which two or more factors each would have been sufficient to produce the injury, so that none of them was a necessary condition for the injury. The various tests that have been proposed as refinements of, supplements to, or substitutes for the but-for test are then examined. All these tests, with the exception of the "NESS" (Necessary Element of a Sufficient Set) test, are shown to be inadequate. The NESS test, which incorporates the traditional Humean philosophic account of the meaning of causation, as modified by John Stuart Mill, was first suggested by Hart and Honoré. Their brief exposition of this test, however, was submerged and weakened by their attempt to treat the tortious-conduct and proximate-cause inquiries as part of the causal inquiry. In Part II the NESS test is more fully elaborated. It is

Theory of Strict Liability, 2 J. LEGAL STUD. 151, 163, 168-69 (1973). Compare Borgo, *supra* note 11, at 425 & n.17, 431 & n.31, 432-40, 452-55 & nn. 55 & 59 (selecting responsible causes) with H.L.A. HART & A. HONORÉ, *supra* note 10, at 26-47, 58-76 (same).

14. For my critique of the accounts of the economic analysts, see Wright, *supra* note 9.

15. Throughout this Article, the arguments presented will be descriptive rather than normative. The normative issues will be addressed in a separate article now in progress.

then used to resolve the problematic causation cases that have resisted solution under all the alternative tests. In the final section of Part II, the steps involved in applying the NESS test are analyzed to demonstrate that, contrary to the currently popular view among legal writers, the causal inquiry is a factual inquiry not dependent on policy considerations.

In Part III I address the difficult issues raised by the "increased risk," "reduced chance," and "alternative liability" cases—cases in which it can be proven that the defendant tortiously exposed the plaintiff to an increased risk of injury but not that the defendant actually contributed to the subsequent injury. It is commonly argued that the causation requirement is properly ignored or relaxed in these cases, or that the requirement should be satisfied by a probabilistic "increased risk" concept of "causation." I contend that these cases are more plausibly and fruitfully viewed as cases that recognize a new type of injury—risk exposure—in certain narrowly circumscribed situations. Causation of this risk-exposure injury is easily established using the basic concept of causation embodied in the NESS test. Finally, I argue that the probabilistic increased-risk concept of causation must be rejected, because it lacks the attributive element that distinguishes causal explanations from mere probability statements. This attributive element, which has always been essential for tort liability, explains the courts' refusal to admit pure or "naked" statistical evidence as proof of causation or identification.

I

THE SCOPE AND SIGNIFICANCE OF THE CAUSAL INQUIRY IN TORT LAW

A. *The Scope of the Causal Inquiry: Distinguishing the Causal Inquiry from the Tortious-Conduct and Proximate-Cause Inquiries*

A number of legal scholars have equated causation of injury with legal responsibility for the injury.¹⁶ Relying on this assumed equivalence, some of these writers argue that because the determination of responsibility is policy-dependent, the causal inquiry also must be policy-dependent. Conversely, others argue that because the causal inquiry is policy-free, the responsibility determination also must be policy-free.

In this Section it will be demonstrated that both of these competing arguments are built on a false premise. Causation is *not* equivalent to responsibility. Rather, before a decision is made on legal responsibility,

16. See *supra* notes 12 & 13 and accompanying text; see also Green, *supra* note 8, at 545, 562; Malone, *supra* note 6, at 62-67; cf. RESTATEMENT (SECOND) OF TORTS § 431 comment a (1965) (substantial factor test of causation includes notion of responsibility).

the causal inquiry must be combined with the tortious-conduct inquiry—"Was the defendant's conduct tortious?"—and the proximate-cause inquiry—"Is there an applicable policy or principle which absolves the defendant from liability, even though his tortious conduct was a cause of the injury?"

As we shall see, the writers who equate causation with responsibility erroneously attribute to the causal inquiry issues that are properly associated with the tortious-conduct and proximate-cause inquiries. Three different examples will be discussed. The first is drawn from Wex Malone's influential article on causation in tort law. In this article, Malone contends that the causation requirement is just one more policy-dependent doctrine that is manipulated by the judge or jury to reach a desired result. The second and third examples are Hart and Honoré's and Richard Epstein's respective theories of tort liability, which attempt to avoid ad hoc, policy-dependent liability determinations by basing liability on expanded notions of causation.

1. *Malone*

Malone's article contains a series of arguments that are meant to establish the policy-dependent nature of the causal inquiry. Only one of them is pertinent here.¹⁷

This argument focuses on the policy considerations that underlie any decision to identify one of several contributing factors as "the cause" of an injury. Malone notes that, depending on their respective backgrounds and motivations, different persons may identify different contributing factors as the cause of an injury. In one of his examples, an irresponsible youth drives too fast down a road recently covered with loose gravel, and a stone is thrown by a wheel of the car into the face of a pedestrian. Malone observes that the neighbors may say that the parents' letting their irresponsible son drive was the cause of the harm, while a road engineer may say that the cause was improper road construction, and a physics teacher may say that the impact of the wheel's momentum on the rock was the cause. Thus, Malone states, causal determinations are evaluative and purposive.¹⁸

In Malone's other example, an elderly worker with a long-standing heart ailment dies of heart failure while engaged in some trivial task for his employer. A medical expert, says Malone, is likely to testify that the trivial task was not a cause of the death, since he will view it as an instance of a commonplace recurrent event with which medical science is helpless to deal. But a judge, considering the compensatory purposes of

17. Malone's other arguments will be discussed later. See *infra* text accompanying notes 240-42 & 304-30.

18. Malone, *supra* note 6, at 62.

the workers' compensation statute, might decide, or allow a jury to decide, that the trivial exertion was a work-related cause of the elderly employee's death.¹⁹ Malone concludes:

"[S]imple" causation is not merely an abstract issue of fact. . . . [T]he resolution of the cause problem depends largely upon the purpose for which cause is to be used. . . . It is through the process of selecting what is to be regarded as a cause for the purpose of resolving a legal dispute that considerations of policy exert their influence in deciding an issue of cause-in-fact.²⁰

This argument confuses several distinct parts of the liability analysis. It is generally agreed that the purpose and context of the particular inquiry determine which of several contributing factors will be selected as "the cause."²¹ But many writers, including Malone, do not recognize that the influence of purpose or policy extends only to the noncausal parts of this selection process: the focus on only one or a few of the many potential causes, and the limitation of responsibility to only certain types of consequences.

Thus, in the elderly worker case, policy considerations are quite important in deciding which causes and consequences the workers' compensation statute was meant to cover. At a minimum, such statutes require that the employee's work or working conditions have contributed to (been a cause of) the injury. That is, only work-related activities and conditions are relevant as potential causes. This is a policy decision. The trivial exertion was a work-related exertion. But did it contribute to the failure of the employee's heart? That is the causal question. The causal issue need not be addressed if it is decided that an injury will not be treated as work-related, even if the injury occurred while at work, when—as seems to be the case here—the risk of such injury was not increased by being at work and, moreover, the risk was personal to the employee. This also is a policy issue.²²

If the court decides that such injuries are work-related, the causal issue must be addressed. The judge should clarify the causal issue for the medical expert by explaining that the medical controllability of the risk is irrelevant. The question is rather whether the exertion in fact contributed to the employee's death by triggering or accelerating the failure of

19. *Id.* at 63-64.

20. *Id.* at 64.

21. *E.g.*, T. BEAUCHAMP & A. ROSENBERG, HUME AND THE PROBLEM OF CAUSATION 261-62, 282, 284-95 (1981); H.L.A. HART & T. HONORÉ, *supra* note 8, at lxxvi, 11-12, 35-38, 62; Borgo, *supra* note 11, at 430-32, 439-40; Cohen, *Field Theory and Judicial Logic*, 59 YALE L.J. 238, 251-59 (1950); Cole, *supra* note 7, at 459-60, 462-65; Mackie, *Causes and Conditions*, in CAUSATION AND CONDITIONALS 21-24 & n.15 (E. Sosa ed. 1975); Williams, *Causation in the Law*, 1961 CAMBRIDGE L.J. 62, 63-65, 69.

22. See EPSTEIN, TORTS, *supra* note 4, at 925-27; 1 A. LARSON, THE LAW OF WORKMEN'S COMPENSATION § 7.20 (1984). See generally *id.* §§ 6, 7, 12 & 13.

his heart. Malone's causal-selection argument does not demonstrate that policy considerations do or should play any role in this causal part of the analysis. All Malone's argument shows is that policy considerations determine which causes and consequences will give rise to liability.

The elderly worker example is a workers' compensation case, rather than a tort case. But the liability analysis in tort law proceeds in the same manner. First, those potential causes that may give rise to tort liability are identified. We are not interested in all the possible causes, but only those that were tortious. This is the tortious-conduct inquiry. Policy considerations determine whether certain conduct will be treated as tortious.²³

The second step, after the identification of tortious conduct which may have contributed to the injury, is the application of the actual-causation requirement, which requires that the tortious conduct actually have contributed to the injury.²⁴ This is the causal inquiry. At this stage it is irrelevant that there may also be other contributing factors (causes).

In fact, it is never necessary for the court to identify or investigate all the contributing factors. It usually will consider only a few contributing factors other than the defendant's tortious conduct: those which might reduce or eliminate, for reasons of policy or principle, the defendant's legal responsibility for harm that was caused by his tortious conduct.²⁵ This is the proximate-cause inquiry.

All three steps—the tortious-conduct inquiry, the causal inquiry, and the proximate-cause inquiry—are involved in determining whether the defendant's conduct was "the cause" of the injury. The causal inquiry determines whether the defendant's conduct was a *cause* of the injury. The tortious-conduct and proximate-cause inquiries determine whether the defendant should be held legally responsible as "the" cause of the injury. That is, the phrase "the cause" is simply an elliptical way of saying "the *responsible* cause." Malone asserts that the determinations involved in establishing that a certain factor not only was a cause but also should be held responsible as "the" cause often are inseparable in the conversations and minds of ordinary people.²⁶ The distinction may not be explicit in conversations, but it is certainly implicit. It is essential to distinguish each step in the adjudication of legal disputes to ensure that the causal inquiry will focus only on the legally relevant potential causes and not be confused with the noncausal issues, as occurred in Malone's

23. The traditional types of tortious conduct recognized by the courts are intentional, negligent, and ultrahazardous behavior.

24. More precisely, the tortious aspect of the conduct must have contributed to the injury. See *infra* text accompanying notes 100-64.

25. J. FLEMING, *supra* note 4, at 171; Green, *supra* note 8, at 548, 557-59, 564; Williams, *supra* note 21, at 63-65.

26. Malone, *supra* note 6, at 66-67.

discussion of the elderly-worker example.²⁷

2. Hart and Honoré

Like Malone, Hart and Honoré apply the causal label to a combined tortious-conduct, cause-in-fact, and proximate-cause inquiry. Unlike Malone, they argue that the combined inquiry is a factual inquiry based on commonsense causal principles, rather than a policy-dependent inquiry based on ad hoc judgments of legal purpose or social expediency.²⁸ However, their commonsense principles are essentially the negligence and intentional tort prongs of the tortious-conduct inquiry and have nothing to do with causation. The causal label is not only misleading, but also makes it difficult for them to account for strict tort liability and certain types of proximate-cause cases.

Hart and Honoré divide their causal analysis into two steps. The first step is the cause-in-fact inquiry, which determines whether the defendant's conduct actually contributed to the injury. If it did, it was a condition for the occurrence of the injury.²⁹ The second step is a conflation of the tortious-conduct and proximate-cause inquiries, in which Hart and Honoré's commonsense "causal" principles are used to determine whether the defendant's conduct can be distinguished from the other contributing factors ("mere conditions") as "the cause" of the injury.³⁰

According to Hart and Honoré, the central notion in the common-sense concept of causation is that the cause is the factor which "makes a difference" by interfering with, intervening in, or otherwise changing the normal or reasonably expected course of events. Thus, a contributing factor is treated as the cause rather than as a mere condition if it was (1) a voluntary human intervention that was intended to produce the consequence (for example, deliberately breaking a vase) or (2) an abnormal action, event, or condition in the particular context (for example, a freak

27. A. BECHT & F. MILLER, *supra* note 8, at 5-7, 12-13; Williams, *supra* note 21, at 63-65, 69. In another "Rumination" published fourteen years later, but much less well known (outside Louisiana), Malone reversed his position and insisted that the causal inquiry is purely factual and should be sharply distinguished from the independent policy considerations that enter into determination of the responsible cause. Malone, *Ruminations on Dixie Drive It Yourself Versus American Beverage Company*, 30 LA. L. REV. 363, 370-71 (1970). For a valiant effort to reconcile the two Ruminations, see Note, *When Cause-in-Fact Is More Than a Fact: The Malone-Green Debate on the Role of Policy in Determining Factual Causation in Tort Law*, 44 LA. L. REV. 1519, 1540-41 & n.89 (1984).

28. See *supra* note 11.

29. H.L.A. HART & T. HONORÉ, *supra* note 8, at 109-11; *see id.* at 72. Hart and Honoré use the "necessary element of a sufficient set" test in this cause-in-fact inquiry. *See infra* text accompanying notes 226-83.

30. H.L.A. HART & T. HONORÉ, *supra* note 8, at 1-2, 24-25, 33, 72-73, 110-11.

storm or driving at an excessive speed).³¹ If there are two or more contributing factors which satisfy one of these criteria, the last one to occur is treated as the cause. When searching for the cause of an injury, we do not trace back any further once we come across a deliberate intervention or an independent abnormal condition.³²

It soon becomes clear that these criteria are neither policy-neutral nor causal. As applied to human conduct in a tort case, the two criteria simply call for inquiries into the intentional or negligent character of the conduct. This is, in effect, a restricted form of the tortious-conduct inquiry.

Thus, a human intervention is considered to be voluntary only if it was a free, deliberate, and informed act or omission, intended to bring about what in fact happened in the manner in which it happened. An intervention is not voluntary if the consequence was not intended, or even if it was intended if the intervention occurred as a result of mistake, ignorance, coercion, or the pressure of moral or legal obligation, or if the choice to intervene was the lesser of two evils.³³ More generally, Hart and Honoré admit that their voluntary-intervention criterion calls for

31. *Id.* at 2, 29, 33-35, 41-42, 130-31. Both editions of Hart and Honoré's book use these criteria as the bases for the initial distinction between causes and "mere conditions." *See id.*; H.L.A. HART & A. HONORÉ, *supra* note 10, at 27, 31-33, 38-39, 130-31, 131 n.1. This part of their analysis corresponds to the tortious-conduct inquiry and is the primary focus of my critique. The second part of their analysis—the proximate-cause inquiry—is concerned with the problems that arise when two or more contributing conditions satisfy one of the two criteria. Generally, they treat the last deliberate intervention or independent abnormal condition as the cause of the injury. Any prior conditions which also satisfy one of the criteria are reduced to the status of "mere conditions." *See infra* note 32.

In the first edition, Hart and Honoré use the same notion of voluntary intervention in both the proximate-cause inquiry and the tortious-conduct inquiry. H.L.A. HART & A. HONORÉ, *supra* note 10, at 38-39, 129-31. In the second edition, however, they use a different definition in the proximate-cause inquiry. They replace the requirement that the intervenor have intended the actual injury with a requirement that he have intended to exploit the situation created by the defendant. H.L.A. HART & T. HONORÉ, *supra* note 8, at 136 & n.23. This definition would not work in the tortious-conduct inquiry, since that inquiry evaluates the conduct of the defendant herself. Moreover, it also appears to be unworkable in the proximate-cause context, unless it is interpreted to require that the intervenor be aware of the "untoward" risks involved in the situation created by the defendant and that the intervenor deliberately (tortiously) act to exploit those risks. *See id.* at 136. Otherwise, any voluntary action affecting the situation created by the defendant would be a superseding cause.

In both editions, the notion of abnormal conditions employed in the proximate-cause inquiry is different than the one employed in the tortious-conduct inquiry. In the proximate-cause inquiry, the notion of abnormality is expanded to include not only conditions which are *per se* unusual but also unusual or coincidental conjunctions of normal events. Conversely, Hart and Honoré state that intervening human actions do not relieve the defendant of liability unless they are grossly abnormal. *Id.* at 77-81, 162-66, 184-85; H.L.A. HART & A. HONORÉ, *supra* note 10, at 151-55, 169-70.

The problems created by Hart and Honoré's attempt to fit the tortious-conduct and proximate-cause inquiries into a unitary causal rubric are discussed further *infra*, at text accompanying notes 42-47.

32. H.L.A. HART & T. HONORÉ, *supra* note 8, at xlviii-xlix, lxxviii-lxxix, 42-43, 49-50, 71, 74, 77-80, 130-31, 135-36, 162-63.

33. *Id.* at 41-42, 75-77, 138 & n.40; *see supra* note 31.

policy judgments on matters of degree and reasonableness. Many of the circumstances which are treated as rendering even deliberate conduct nonvoluntary correspond to the recognized moral and legal grounds for justifying or excusing harmful behavior. Examples include reasonable actions taken to defend persons or property against reasonably perceived perils, to safeguard legal rights or interests, to rescue others, and so forth.³⁴ Overall, Hart and Honoré's voluntary-intervention criterion simply describes one of the principal recognized types of tortious conduct: the intentional, unjustified, and unexcused infliction of injury.³⁵

Similarly, the abnormal-condition criterion also calls for nonfactual and noncausal judgments. When it is applied to human action, it is essentially identical to the negligence category of tortious conduct, particularly when attribution of responsibility is at issue. Hart and Honoré note that the distinction between the abnormal condition and the other normal conditions is not based on differences in actual contributions to the consequence, but rather depends, as Malone pointed out, on the context and purpose of the inquiry.³⁶ They state that abnormal human conduct is conduct that deviates from the usual, expected, or established standards of behavior. They treat omissions as abnormal conditions only when the failure to act constitutes a breach of duty—that is, when the omission is negligent.³⁷ More generally, they observe that, “[i]n relation to human conduct, . . . the notion of what is ‘natural’ is strongly influenced by moral and legal standards of proper conduct, though weight is also given to the fact that certain conduct is usual or ordinary for a human being.”³⁸

In sum, as Hart and Honoré acknowledge,³⁹ their two “causal” criteria (voluntary or abnormal conduct) do not have anything to do with the inquiry into actual contribution to the injury, which is the causal aspect of responsibility denoted by the word “cause” in the phrase “the cause.” Instead, the two criteria are relevant only in determining whether the defendant's conduct was “the” (responsible) cause. The two criteria accomplish this task by focusing on the tortious (intentional or negligent) character of the conduct that contributed to the injury.

Thus, it is the combination of causation (the causal inquiry), tor-

34. H.L.A. HART & T. HONORÉ, *supra* note 8, at 1v, 138, 141-60.

35. However, the criterion is too narrow to serve as a complete description of tortious intentional conduct. *See infra* text accompanying notes 45-47.

36. H.L.A. HART & T. HONORÉ, *supra* note 8, at lxxvi, 11-12, 35-38, 62, 72, 110-11.

37. *Id.* at 37-38, 50-51, 59-60, 64, 135 n.15, 138-41, 183-85.

38. *Id.* at 183. The correspondence between the abnormal conduct criterion and negligence inspired Abraham Harari's effort to demonstrate that negligence (as a supposed aspect of causation) is the basis of all tort liability. A. HARARI, *supra* note 13, at 49-55. One of the weaknesses in Harari's argument is the assumption that the abnormal conduct criterion is a causal criterion.

39. H.L.A. HART & T. HONORÉ, *supra* note 8, at xlvi-xlix, 72-74, 110-11.

tious conduct by the defendant (the tortious-conduct inquiry), and the absence of any intervening abnormal physical occurrences or tortious behavior by third parties (the proximate-cause inquiry) that creates the strong and usually sufficient case for legal responsibility. Hart and Honoré's statements treating causation as the primary ground of legal responsibility are plausible only when it is understood that they are using the term "causation" to refer to this combination of the three different elements of tort liability analysis.⁴⁰ They nevertheless insist that their approach is based almost exclusively on causal principles. This insistence seems to be motivated by a desire to distinguish their common-sense, principled account of proximate-cause analysis from the social-policy accounts of the legal realists and others.⁴¹ The causal label creates an appearance of policy-neutrality, which contrasts with the explicit and broader notion of policy in the alternative accounts.

However, the causal label is misleading and unnecessary, and it creates major difficulties for Hart and Honoré.

For example, since Hart and Honoré insist that the defendant's conduct was a cause of an injury only if it was deliberate or abnormal (negligent), they are forced to classify strict liability as a noncausal form of liability.⁴²

Similarly, because they treat the *last* deliberate intervention or independent abnormal occurrence as the cause of an injury, to the exclusion of all prior interventions or occurrences,⁴³ they find it difficult to account for cases in which liability is imposed on the defendant despite some subsequent deliberate intervention or independent abnormal occurrence. They describe these as cases of "inducing" or "occasioning" harm rather than "causing" harm. They claim that there is a "causal connection" only in a metaphorical sense, even though the defendant's tortious conduct clearly contributed to the injury. They assert that responsibility in these cases is based on noncausal policy considerations, and they therefore use tortious conduct language rather than their usual causal language to state the grounds for (and limits on) liability. The defendant is held liable only if he *intentionally* induced the subsequent deliberate action of another or *negligently* provided the opportunity for a subsequent deliberate intervention or abnormal occurrence.⁴⁴

40. See *supra* note 12 and accompanying text; see also H.L.A. HART & T. HONORÉ, *supra* note 8, at 117-18, 131-85, 135 n.9, 205-53 (tort liability discussed in terms of causal connection between tortious conduct and harm, and liability negated despite such causal connection if there was an intervening abnormal physical occurrence or intervening tortious behavior).

41. H.L.A. HART & T. HONORÉ, *supra* note 8, at xxxv-xxxvii, xlvi-lv, 3-7, 24-25, 65-67, 88-111, 130-32, 254, 304-07.

42. See *id.* at 85-86.

43. See *supra* note 32.

44. H.L.A. HART & T. HONORÉ, *supra* note 8, at 51-52, 57, 59-61, 71, 81-83, 133, 186. Hart

Finally, Hart and Honoré face an insuperable dilemma when they attempt to formulate definitions of their voluntary-intervention and abnormal-condition criteria that will properly handle both the tortious-conduct inquiry and the proximate cause inquiry. Since these are supposedly causal criteria, their content should not vary when applied to the tortious-conduct inquiry as opposed to the proximate-cause inquiry. Yet, in practice, broader definitions of "voluntary" and "abnormal" are required for the tortious-conduct inquiry than for the proximate-cause inquiry. In the proximate-cause inquiry, narrow definitions are needed to avoid cutting off the defendant's liability when subsequent contributing conduct, although intentional ("voluntary"), was justified or done in ignorance of the circumstances or the possible consequences or, although negligent ("abnormal"), was "routine" rather than "highly unusual."⁴⁵ In the tortious-conduct inquiry, on the other hand, broader definitions are needed. The defendant's intentional ("voluntary") intervention often may be treated as tortious despite ignorance, mistake or necessity, and his negligent ("abnormal") conduct will be treated as tortious whether it is "routine" or "highly unusual."⁴⁶ Hart and Honoré respond to this dilemma by adopting a definition of "voluntary" that is suitable for the proximate-cause inquiry but too narrow for the tortious-conduct inquiry and, conversely, a definition of "abnormal" that is suitable for the tortious-conduct inquiry but too broad for the proximate-cause inquiry.⁴⁷

Hart and Honoré could have avoided these and other difficulties easily had they correctly described their commonsense principles as principles of responsibility rather than as principles of causation, and had they separated the proximate-cause element in these principles from the tortious-conduct and causal elements. As with Malone's argument, Hart

and Honoré's position on this issue is influenced by their view that volitional human actions, unlike physical events, are neither subject to nor explainable in terms of causal generalizations. *Id.* at 51-52, 55-57, 60-61; *accord* Kadish, *Complicity, Cause and Blame: A Study of the Interpretation of Doctrine*, 73 CALIF. L. REV. 323, 326-27, 332-35 & n.8 (1985); Williams, *supra* note 21, at 66-68. But, with human actions just as with physical events, if *all* the relevant circumstances (accumulated experience and knowledge, mood, and so forth) were the same, surely the decision or action also would be the same. To assert otherwise is to assert that human action is random or arbitrary. Cf. H.L.A. HART & T. HONORÉ, *supra* note 8, at 56-57 (acknowledging that generalizations apply to human action, but nevertheless denying repeatability in identical circumstances). Human action is less regular and predictable than physical events because humans learn from prior experiences and new information, because the range of relevant conditions is much broader, and because the applicable causal generalizations are much more complex and less well understood. See generally T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 314-27; J. MACKIE, *THE CEMENT OF THE UNIVERSE* 120-26 (1974); Moore, *Causation and the Excuses*, 73 CALIF. L. REV. 1091, 1112, 1124-27, 1132-37 (1985).

45. H.L.A. HART & T. HONORÉ, *supra* note 8, at 41-42, 75-77, 136, 152-53, 183-84; W. PROSSER, *supra* note 1, § 44, at 274, 276-78, 282.

46. W. PROSSER, *supra* note 1, §§ 7-8, at 30-33; *id.* § 9, at 35; *id.* § 17, at 99-100; *id.* § 24, at 126-27; *id.* § 31, at 145; *id.* § 32, at 150-51.

47. See *supra* note 31 and text accompanying notes 33-38.

and Honoré's theory illustrates the conceptual confusion and practical difficulties that inevitably result when these three elements of tort liability analysis are not carefully distinguished from one another.

3. Epstein

Hart and Honoré's book has inspired several other attempts to develop causal theories of tort liability.⁴⁸ The best known example in the United States is Richard Epstein's theory of "strict" liability.⁴⁹

Epstein, even more than Hart and Honoré, opposes the ad hoc, social-policy accounts of tort liability, including the currently fashionable wealth-maximization accounts.⁵⁰ He argues that tort liability should be based on principles of individual autonomy and responsibility, with the dividing line between autonomy and responsibility being marked by causation of harm to others.⁵¹ Epstein's emphasis on individual autonomy is coupled with a strong preference for precise rules of responsibility. He seeks to eliminate or greatly restrict those aspects of traditional tort liability analysis—such as tests of reasonableness, negligence, or proximate cause—which, because of their imprecision, can be used by judges or juries to validate substantial incursions on individual autonomy.⁵²

Although Epstein refers to his theory as a theory of strict liability, it is in fact a system of defeasible absolute liability. Tort theories are classified as "absolute," "strict," or "fault-based" depending on the elements required to state a minimally sufficient *prima facie* case.⁵³ Traditionally, absolute liability has meant *prima facie* liability based merely on causa-

48. See *supra* note 13. Abraham Harari's causal theory is discussed briefly *supra* in note 38. John Borgo's analogical reasoning approach suffers from precisely the same flaw that Borgo correctly identifies in Epstein's approach. It abstracts responsibility paradigms from situations while overlooking critical aspects of those situations and then tries to apply the abstracted paradigms to situations where the critical aspects are different. See Borgo, *supra* note 11, at 429-31, 436-40.

49. Epstein, *supra* note 13; Epstein, *Defenses and Subsequent Pleas in a System of Strict Liability*, 3 J. LEGAL STUD. 165 (1974) [hereinafter cited as Epstein, *Defenses*]; Epstein, *Intentional Harms*, 4 J. LEGAL STUD. 391 (1975) [hereinafter cited as Epstein, *Intentional Harms*]; Epstein, *Nuisance Law: Corrective Justice and its Utilitarian Constraints*, 8 J. LEGAL STUD. 49 (1979) [hereinafter cited as Epstein, *Nuisance Law*]; Epstein, *Causation and Corrective Justice: A Reply to Two Critics*, 8 J. LEGAL STUD. 477 (1979) [hereinafter cited as Epstein, *Reply*]. Hart and Honoré's influence on Epstein is indicated in, for example, Epstein, *supra* note 13, at 161-64 & n.39; Epstein, *Intentional Harms*, *supra*, at 431-32 & n.100; Epstein, *Reply*, *supra*, at 479, 486 n.33. Compare Epstein, *supra* note 13, at 166-69, 175 (building on simple transitive propositions), with H.L.A. HART & A. HONORÉ, *supra* note 10, at 26-28, 68-69 (same). However, in their second edition Hart and Honoré reject Epstein's "causal maximalism." H.L.A. HART & T. HONORÉ, *supra* note 8, at lxxiii-lxxvii; see *id.* at 30-31, 96-97 (notion of force is inadequate basis for causal theory).

50. Epstein, *supra* note 13, at 151-57, 198-99.

51. *Id.* at 198-201, 203-04; Epstein, *Intentional Harms*, *supra* note 49, at 441-42; Epstein, *Reply*, *supra* note 49, at 479-80.

52. Epstein, *supra* note 13, at 198-99; *see id.* at 162-66, 185-87; Epstein, *Defenses*, *supra* note 49, at 178-81, 184; Epstein, *Reply*, *supra* note 49, at 478-80 & n.5.

53. See Epstein, *Pleadings and Presumptions*, 40 U. CHI. L. REV. 556, 558-59, 570-71, 577 (1973) (minimally sufficient elements for responsibility constitute *prima facie* case and establish basic

tion of harm to another.⁵⁴ Strict liability requires, in addition to causation, that the harm result from conduct which created a foreseeable, unaccepted risk of harm to another that *could* have been avoided.⁵⁵ Fault liability requires, in addition to causation, that the harm result from conduct which *should* have been avoided given the foreseeable risks.⁵⁶ Epstein purports to base *prima facie* liability simply on causation of harm to another.⁵⁷ This *prima facie* absolute liability, however, may be eliminated or reduced if certain defenses are established.⁵⁸

Epstein's normative arguments on behalf of his absolute liability theory assume that there are only two possible types of *prima facie* tort liability—fault and absolute.⁵⁹ His negative arguments point out the flaws in the fault theories and the failure of those theories to account for various cases in which liability was imposed (or arguably should have been imposed) although there was no faulty behavior by the defendant.⁶⁰ The cases he discusses, however, are strict liability cases—cases in which the defendant knew or should have known that he was imposing an unaccepted risk on others.⁶¹ Similarly, his affirmative arguments are almost

theory of recovery, even when responsibility defeasible through defenses or other subsequent pleadings.

54. E.g., O.W. HOLMES, *supra* note 1, at 67-68, 72; Ehrenzweig, *Negligence Without Fault*, 54 CALIF. L. REV. 1422, 1426 nn.13 & 18 (1966), and the sources cited therein.

55. E.g., PROSSER & KEETON, *supra* note 7, § 75, at 536-38; *id.* § 78, at 555-56; Ehrenzweig, *supra* note 54, at 1450-51 & n.1; see O.W. HOLMES, *supra* note 1, at 77, 93-95, 115-18.

56. E.g., O.W. HOLMES, *supra* note 1, at 85-88; Ehrenzweig, *supra* note 54, at 1444-45.

57. Epstein, *supra* note 13, at 152, 168-69.

58. These defenses include reciprocal causation of harm to the defendant by the plaintiff, assumption of the risk by the plaintiff, and trespass on the defendant's property by the plaintiff. Epstein, *Defenses*, *supra* note 49, at 167-68, 174, 185, 201.

59. Epstein, *supra* note 13, at 152, 171; Epstein, *Intentional Harms*, *supra* note 49, at 398. This erroneous assumption also appears in most economic analyses of tort liability. See, e.g., Brown, *Toward An Economic Theory of Liability*, 2 J. LEGAL STUD. 323 (1973); Shavell, *Strict Liability Versus Negligence*, 9 J. LEGAL STUD. 1 (1980).

60. Epstein, *supra* note 13, at 153-60, 169-71.

61. Morris v. Platt, 32 Conn. 75 (1864); Vincent v. Lake Erie Transp. Co., 109 Minn. 456, 124 N.W. 221 (1910). Both cases are discussed in Epstein, *supra* note 13, at 157-60. See also Bolton v. Stone, 1951 A.C. 850, *rev'd* [1950] 1 K.B. 201 (C.A. 1949), *rev'd* [1949] 1 All. E.R. 237 (Manchester Assizes 1948). *Bolton* is discussed in Epstein, *supra* note 13, at 169-71. In *Bolton*, the defendant cricket club maintained and used a playing field not only with the knowledge that cricket balls might be hit out of the grounds and endanger passersby such as the plaintiff, but also with the knowledge that one of the objects of the game is to hit the ball away from the opposing team, the further away the better. Although balls are rarely hit out due to the size of the playing fields, those that are hit out earn extra points. See 1 THE ENCYCLOPEDIA OF SPORT 447 (1898). Thus, although the risks to passersby such as the plaintiff were slight, the possibility was knowingly and eagerly pursued as an integral part of the defendant's activity. This latter aspect of the case, together with the likely severity of the harm should a ball actually escape the grounds and hit someone, may well explain (a) the House of Lords' unease in refusing to hold the defendant liable, (b) the public interest in the case, and (c) the decision of the Cricket Clubs of England to compensate the plaintiff despite the lack of a holding of legal liability. See Epstein, *supra* note 13, at 170. It seems unlikely that the case would have aroused such controversy if the possibility that cricket balls would escape the grounds had been completely unintended, unknown, and unforeseeable.

always couched in strict-liability terms. They explicitly or implicitly rely not only on causation of harm by the defendant, but also on the defendant's acting despite actual or constructive knowledge of a foreseeable, unaccepted risk to another.⁶²

This mental element, however, is absent in Epstein's description of the *prima facie* case under his theory, which initially is based solely on causation of harm. Epstein does not provide any general definition or account of causation. He rejects the traditional "but for" (necessary condition) test because it treats too many conditions as causes of the injury and therefore allegedly throws all the critical liability decisions into the policy-laden proximate-cause inquiry.⁶³ In place of the but-for test and the proximate-cause inquiry, he identifies a set of "causal paradigms" which purportedly explicate the concept of causation and constitute necessary and *prima facie* sufficient bases for the imposition of liability.⁶⁴

But neither of these claims holds up. Epstein's paradigms do not explicate the concept of causation.⁶⁵ Instead, he uses the concept of causation, undefined and unelaborated, to construct what are actually paradigms of responsibility, by combining certain methods of causing harm (primarily force and the threat of force) with certain noncausal criteria. However, the prescribed methods of causing harm are too restrictive, while the prescribed noncausal criteria do not include the mental element required for true strict liability. Consequently, Epstein's theory produces results which diverge substantially from those reached by the courts.⁶⁶

The first paradigm is "*A* hit *B*," which is meant to be an abbreviated way of saying that *A*, by application of force to *B* or *B*'s property, broke, cut, or otherwise harmed *B* or *B*'s property.⁶⁷ Note the implicit circular reference to the concept of causation in this "causal" paradigm. The

62. Epstein, *supra* note 13, at 158-59 & n.22, 173; Epstein, *Defenses*, *supra* note 49, at 169, 213; Epstein, *Intentional Harms*, *supra* note 49, at 398, 414.

63. Epstein, *supra* note 13, at 161-63; Epstein, *Reply*, *supra* note 49, at 478-79. But see Epstein, *supra* note 13, at 183 (but-for test applied to dangerous position of vase), 184 (same: dangerous condition of car).

64. Epstein, *supra* note 13, at 160, 165-66, 168-69, 190-91, 194-95; Epstein, *Defenses*, *supra* note 49, at 167-68; Epstein, *Intentional Harms*, *supra* note 49, at 398-400; Epstein, *Reply*, *supra* note 49, at 479-80.

65. This point has also been made by Borgo, *supra* note 11, at 427-31; see *supra* note 48.

66. Epstein states that his theory is predominantly normative. Epstein, *supra* note 13, at 151; Epstein, *Reply*, *supra* note 49, at 477. Nevertheless, he recognizes that any plausible normative theory cannot stray too often or too far in its results from the commonly accepted notions of responsibility embodied in traditional tort law, and he claims to base his theory on those commonly accepted notions. See Epstein, *supra* note 13, at 151, 163-64, 166; Epstein, *Reply*, *supra*, at 479, 503-04. Indeed, he argues that his theory best accounts for the results reached by the courts. Epstein, *supra* note 13, at 165-66, 189; Epstein, *Defenses*, *supra* note 49, at 168-69, 173-74, 204-05, 213-15; Epstein, *Intentional Harms*, *supra* note 49, at 406-07; Epstein, *Reply*, *supra*, at 480.

67. Epstein, *supra* note 13, at 166-67; Epstein, *Intentional Harms*, *supra* note 49, at 399; Epstein, *Reply*, *supra* note 49, at 480-81.

paradigm implicitly states that *A*, by application of force, *caused* harm to *B* or *B*'s property—for example, that the force initiated (caused) by *A* caused *B*'s cut, which is itself the harm for which compensation is sought or which led to (caused) the ultimate harm for which compensation is sought. The paradigm does not explain or help to identify any of these instances of causation. It rather assumes that we somehow perceive them, and it restricts liability to those that involve force as the causal mechanism.

Epstein also requires that the force have been initiated by a volitional act of *A*, rather than by someone's throwing *A* or by some automatic or reflex action of *A*.⁶⁸ Thus the paradigm includes, in addition to (unelaborated) actual causation, the noncausal criteria of a volitional act and a particular method of causation (the application of force).

Similarly, the second paradigm, "*A* frightened *B*," requires, in addition to actual causation, that *A* have caused the harm to *B* by a volitional act and through the causal mechanism of fright or shock. As originally stated, the paradigm would cover cases in which extrasensitive or even normal people are frightened by the completely innocent and inoffensive acts of others. Epstein suggests that, in these cases, the frightened person should be treated as having caused her own fright, but he does not indicate which of his causal paradigms would apply.⁶⁹ None would, especially since the person's fright is not considered to be volitional. Eventually Epstein modifies the paradigm by requiring that *A*'s volitional act be an "offer of force" against *B*, or, to the same effect, that a de minimus rule be applied to *A*'s conduct.⁷⁰ These modifications, of course, introduce additional noncausal criteria into the paradigm.

The third paradigm holds *A* liable to *C* if "*A* [by a volitional act] compelled *B* to hit *C*."⁷¹ The compulsion must be accomplished through force or the threat of force.⁷² Under this third paradigm, *B* as well as *A* is liable to *C*, since *B* satisfies the initial paradigm, "*B* hit *C*." Moreover, *B* has an action against *A* for the damages that he is required to pay *C*, since *A* satisfies the paradigm, "*A* hit (or offered to hit) *B*."⁷³

The fourth paradigm most clearly demonstrates the circularity and ambiguity of Epstein's causal paradigms. It holds *A* liable to *B* if "*A* [by a volitional act] created [caused] the dangerous condition that resulted in

68. Epstein, *supra* note 13, at 166-67. See generally Moore, *supra* note 44, at 1132-36 (distinguishing action from mere causal sequence).

69. Epstein, *supra* note 13, at 172.

70. Epstein, *Reply*, *supra* note 49, at 483-84.

71. Epstein, *supra* note 13, at 174. See generally Moore, *supra* note 44, at 1129-32 (distinguishing compulsion from mere causation).

72. See Epstein, *Defenses*, *supra* note 49, at 175; Epstein, *Intentional Harms*, *supra* note 49, at 399.

73. Epstein, *supra* note 13, at 175; Epstein, *Defenses*, *supra* note 49, at 174-75.

[caused] harm to *B*.⁷⁴ Note that two instances of causation are implied by the expression “resulted in.” First, some act or event must “trigger” the dangerous condition, and, second, the triggered condition must then cause the harm.

Epstein attempts to avoid the charge that this paradigm, by use of the expression “resulted in” (he overlooks the verb “created”), defines causation in terms of itself. He says that the expression “resulted in” must be replaced in each particular case with a description of the act or event which is the immediate cause of the harm, using one of the three previously elaborated paradigms—force, fright or compulsion.⁷⁵ However, even if both the expression “resulted in” and the verb “created” were interpreted and used in this restricted fashion, it would only push the circularity one level deeper, since—as we have seen—the three previous paradigms themselves contain implicit, unelaborated causal statements. Moreover, Epstein does not actually restrict either term in this fashion.

Thus, in Epstein’s examples dangerous conditions are “created” by “storing” or “handling” explosives, by “placing” or “leaving” an object in a precarious or obstructive position, or by “manufacturing” a defective product. And dangerous conditions are “triggered” by a “small change in conditions, like temperature and humidity,” or by mere use of a defective product (which might simply be turning it on or eating it).⁷⁶ Clearly, given these examples, the creation or triggering of dangerous conditions is not limited to the use or threat of force, unless “force” is construed so broadly as to encompass any action.

Similarly, in Epstein’s examples dangerous conditions, when triggered, can cause harm not just by “releas[ing] or otherwise set[ting] in motion large forces” that are associated with the dangerous condition itself, but also by absorbing or “redirecting” forces that are not associated with the dangerous condition. For example, an obstruction in a highway is treated as a dangerous condition that “results in” injury to *B* when *B* drives into it or swerves to avoid it.⁷⁷ Here “results in” cannot be replaced by one of the three previous paradigms: the obstruction did not hit *B*, or offer to hit *B*, or compel anyone to hit *B*. Rather, the converse is true: the obstruction was hit by *B* or, in the swerve case, “offered” to be hit by *B*!⁷⁸

74. Epstein, *supra* note 13, at 177.

75. *Id.*

76. *Id.* at 177-78.

77. *Id.* at 178, 185.

78. Epstein might argue that *A*’s obstruction compelled *B* to hit *A*’s obstruction. This argument, however, would have to rely on a definition of compulsion that includes more than the actual or threatened use of force by *A*. At most, *A* only threatened to absorb *B*’s force. Moreover, the argument would only establish *B*’s right to be indemnified by *A* for damages that *B* caused to *A*’s

In other examples, the triggered condition "results in" harm through causal mechanisms involving fire, pollution, unwanted oxygen, delays, and so forth, in which there is no use or offer of force by anyone, unless once again "force" is interpreted so broadly that it applies to anything.⁷⁹

In addition to its causal circularity and ambiguity, the dangerous-condition paradigm contains the most obvious noncausal criterion of any of the paradigms: the requirement that the condition which resulted in the harm have been "dangerous." Epstein tries to tie the term "dangerous" into his notion of causation-as-force by limiting the paradigm to those "dangerous conditions that release or otherwise redirect forces in the narrowest sense of causation."⁸⁰ Thus, he asserts, a sharp knife is a "mere" condition rather than a "dangerous" condition, because it contains no store of energy waiting to be released (nor, presumably, does it "redirect" forces).⁸¹ But, as we have already seen, Epstein does not adhere to this limitation. An obstruction, when hit by something, does not release or redirect forces. It absorbs forces, as does *B* in the paradigm "*A* hit *B*." Similarly, if a knife with a sharp edge does not release or redirect forces, then neither does a toy that is defective because it has a sharp edge. Yet these two examples clearly illustrate two of Epstein's three categories of dangerous conditions: things in dangerous positions and defective products. (The third category is inherently dangerous or ultrahazardous things, such as stored explosives.)⁸²

The examples discussed so far indicate the accordion-like quality of Epstein's paradigms and his causal theory as a whole, which alternately expand and contract to fit the particular argument. When Epstein wants to deny liability in a certain case or to emphasize the coherence of his theory, he insists on causation through force, the offer of force, or the creation of a dangerous condition which releases or redirects force.⁸³ But strict adherence to this limitation would require him to deny liability in many types of cases in which liability is imposed by the courts: fires, poisonings, fraud, defamation, emotional distress cases not involving the use of force, and so forth. When discussing these types of cases, Epstein

obstruction (or other object, in the swerve case). There is no embedded paradigm, "*A* hit (or offered to hit) *B*" that would support an action by *B* against *A* for damages suffered by *B* when *B* hit the obstruction. *See supra* text accompanying notes 71-73. Nor, unless we are to twist language, can it be said that *A* compelled *B* to hit himself. *B* hit the obstruction (or other object), not himself.

79. Epstein, *Nuisance Law*, *supra* note 49, at 56-57, 68, 100; Epstein, *Reply*, *supra* note 49, at 486-87.

80. Epstein, *supra* note 13, at 185; *see id.* at 179.

81. *Id.* at 179.

82. *See id.* at 177-78; *see also* Epstein, *Intentional Harms*, *supra* note 49, at 415 (barbed wire as a dangerous condition).

83. Epstein, *supra* note 13, at 180, 185; Epstein, *Intentional Harms*, *supra* note 49, at 399-400, 431; Epstein, *Reply*, *supra* note 49, at 479-81.

usually abandons the force limitation and recognizes any "paradigm" that expresses a "nonreciprocal" (transitive) causal relationship—for example, "*A* poisoned *B*," "*A* segregated *B* where aid could not be given," "*A* misled *B*," "*A* entered *B*'s land," "*A* damaged *B*'s property," or "*A* infringed *B*'s interest."⁸⁴

At this point, the alleged power of the causal paradigms disappears. As Epstein implicitly admits, the inquiry has dissolved into a debate over rights and protected interests.⁸⁵ He belatedly notes that the question of nonreciprocal causation of harm—of who harmed whom—cannot be resolved without a prior determination of the parties' respective rights.⁸⁶ Indeed, he goes so far as to identify the "causal question" with "the question of what legal right of the plaintiff was involved."⁸⁷ Unfortunately, he assumes that a person's rights in her body or property are absolute: theoretically, any invasion creates liability.⁸⁸ As a result, he is forced to introduce "utilitarian constraints" into his system to avoid liability for de minimis harms.⁸⁹

Similar problems appear in Epstein's infrequent discussions of proximate-cause limitations. Although Epstein intended to replace traditional proximate-cause analysis with his causal paradigms,⁹⁰ he never demonstrates how his paradigms accomplish this goal, and it is not clear how they could accomplish it. The four initial paradigms merely require that the injury result from force (or an offer of force) initiated by *A* himself or by someone compelled by *A* or by a dangerous condition created by *A*. The basic proximate cause problems remain. How far is the force

84. Epstein, *supra* note 13, at 195 n.104; Epstein, *Defenses*, *supra* note 49, at 202; Epstein, *Intentional Harms*, *supra* note 49, at 432; Epstein, *Nuisance Law*, *supra* note 49, at 51. In the emotional distress cases not involving the use or threat of force, Epstein adheres to the force limitation and denies liability under his "invasion model." He suggests that liability in these cases is based on "breaches of implied contracts for good social behavior" and that the "contractual mold" is appropriate since these cases "all involve direct personal contact between plaintiff and defendant." Epstein, *Nuisance Law*, *supra*, at 64 n.44. This reasoning, of course, would apply to almost every tort. Epstein's discussion demonstrates the difficulties created by his narrow view of tort theory and his frequent resort to dubious contractual arguments to escape those difficulties. See also Epstein, *Reply*, *supra* note 49, at 481 (drivers on highway are not strangers since they are "bound together by . . . statutory . . . bonds," thus calling for "complicated notions of 'avoidance of harm' which are not captured by my paradigms").

85. Epstein, *Intentional Harms*, *supra* note 49, at 432-41.

86. *Id.* at 432-34, 441; Epstein, *Nuisance Law*, *supra* note 49, at 50-53. Compare Epstein, *supra* note 13, at 164-69 (nonreciprocal causation defines rights), with Epstein, *Nuisance Law*, *supra*, at 58-60 (question of causation depends on prior definition of rights) and Epstein, *Reply*, *supra* note 49, at 498 & n.64 (references to Epstein's gradual recognition of the primacy of rights analysis).

87. Epstein, *Intentional Harms*, *supra* note 49, at 433.

88. See Epstein, *Nuisance Law*, *supra* note 49, at 63-64 & n.40, 82, 92-93; Epstein, *Reply*, *supra* note 49, at 488-500.

89. Epstein, *Nuisance Law*, *supra* note 49, at 50, 74-102; Epstein, *Reply*, *supra* note 49, at 483-84, 488-89, 503.

90. Epstein, *supra* note 13, at 161-63, 164-65, 168, 184-85; Epstein, *Reply*, *supra* note 49, at 478-80.

to be traced? Is *A* liable for harm caused by the dangerous condition despite intervening events?

Epstein compares the case in which *B* is injured when he swerves to avoid *A*'s obstruction with the case in which *B* is injured while traveling back and forth to remove *A*'s obstruction. Epstein claims that the dangerous-condition paradigm distinguishes the two cases. *A* is liable only in the swerve case, because in that case *A*'s obstruction "redirected" *B*'s force, and *B* was injured by the redirected force. But this is true also in the traveling-back-and-forth case. Epstein's further assertion that the "causal connection ends once the road is repaired" is unsupported by his paradigms. Moreover, the logic of the assertion would permit (indeed, require) liability if the injury occurred on the way to the obstruction.⁹¹

Epstein sometimes states that the injury must be produced directly and immediately by the force,⁹² but he provides very little elaboration. In some cases, he is willing to treat intervening events as "means,"⁹³ while in other cases he is willing to trace forces through lengthy, complex causal chains.⁹⁴ On the few occasions that he discusses proximate-cause problems, he frequently relies on traditional proximate-cause arguments.⁹⁵

In sum, Epstein's causal paradigms as initially elaborated are much too narrow to serve as the basis for a comprehensive theory of tort liability. When he expands the content and number of the paradigms to make his theory more plausible, the paradigms become superfluous. There is liability for any volitional act that infringes a legally protected interest, and any method of causation qualifies as an infringement, subject to unspecified but apparently traditional proximate-cause limitations.

However, no matter how many methods of causation are allowed, Epstein's theory still diverges substantially from the decisions of the courts, since his paradigms do not include the critical mental element that is required for liability. A volitional act is not enough. At a minimum, the courts require the plaintiff to establish that the defendant knew or should have known that his act would impose an unaccepted risk on others.⁹⁶ Epstein's paradigms do not include this element (except perhaps the fright paradigm when modified to require an "offer" of force).

91. Epstein, *supra* note 13, at 185.

92. *Id.* at 177, 184-185; Epstein, *Intentional Harms*, *supra* note 49, at 399-400; Epstein, *Nuisance Law*, *supra* note 49, at 56.

93. Epstein, *supra* note 13, at 167-68.

94. Epstein, *Nuisance Law*, *supra* note 49, at 56-57 (nuisance cases); see also Epstein, *Intentional Harms*, *supra* note 49, at 429 (defendant used force to drive away plaintiff's potential customers).

95. See Epstein, *supra* note 13, at 174 n.70, 176 n.73, 185; Epstein, *Defenses*, *supra* note 49, at 183-84; Epstein, *Intentional Harms*, *supra* note 49, at 431-32 & n.100; Epstein, *Nuisance Law*, *supra* note 49, at 56-57; Epstein, *Reply*, *supra* note 49, at 486-87 & n.37.

96. O.W. HOLMES, *supra* note 1, at 76-78; see also *id.* at 92-95, 115-18.

The gap between Epstein's theory and the actual practice of the courts is narrowed somewhat in the reciprocal-causation cases. In these cases, both the plaintiff and the defendant have harmed each other under one of the causal paradigms. The plaintiff's reciprocal causation of harm to the defendant provides a defense, in the second stage of the pleadings, to the plaintiff's *prima facie* case against the defendant, which was established in the first stage. Epstein then allows the plaintiff, in a third stage of the pleadings, to reestablish his claim by proving that the defendant's conduct was tortious.⁹⁷ But by then it is too little and too late. The tortious-conduct inquiry is never reached in the many cases in which the pleadings do not progress beyond the first stage because there is no reciprocal causation. Even when the tortious-conduct inquiry is reached, Epstein refuses to consider one of the principal types of tortious conduct—negligence—because of his strong distrust of imprecise liability principles.⁹⁸ Thus, under Epstein's theory, a driver whose car accidentally breaks down on the highway and who sets out adequate warnings for other drivers is liable to, but cannot recover from, the operator of a truck who negligently crashes into the car. The driver of the car creates a dangerous-condition by blocking the right of way, the truck's crashing into the car merely completes the paradigm, and the truck operator's negligence is (according to Epstein) immaterial.⁹⁹

It should by now be clear that causation is not equivalent to responsibility, nor is it *prima facie* sufficient for liability. The tortious-conduct and proximate-cause inquiries are distinct from the causal inquiry, and all three are required to complete the liability analysis under traditional notions of legal responsibility.

An important question remains, however. Is the causal inquiry a significant part of tort analysis, or is it, as alleged by Malone and feared by Hart, Honoré, and Epstein, dominated by the tortious-conduct and proximate-cause inquiries?

97. Epstein, *Defenses*, *supra* note 49, at 177-84; Epstein, *Intentional Harms*, *supra* note 49, *passim*.

98. Epstein, *Defenses*, *supra* note 49, at 172; Epstein, *Reply*, *supra* note 49, at 482; *see supra* note 52 and accompanying text. *But cf.* Epstein, *Defenses*, *supra*, at 177-84 & n.49 (statutory duties); Epstein, *Nuisance Law*, *supra* note 49, at 71-72 & n.58 (reasonableness considerations in highway cases); Schwartz, *The Vitality of Negligence and the Ethics of Strict Liability*, 15 G.A. L. REV. 963, 986-89, 994 (1981) (discussing Epstein's reliance on negligence notions in certain situations); *supra* note 84 (reasonableness considerations in emotional distress cases).

99. Epstein, *supra* note 13, at 180-81, 191-92 & n.99; Epstein, *Defenses*, *supra* note 49, at 177-78.

B. The Significance of the Causal Inquiry: Linking the Causal Inquiry and the Tortious-Conduct Inquiry Through the Tortious-Aspect Causation Requirement

The significance of the causal inquiry in tort law varies dramatically depending on how it is linked to the tortious-conduct inquiry. There are two principal options. The first is to apply the causal inquiry to the defendant's conduct as a whole, for example, operating a hotel or driving a car. The second is to focus the causal inquiry on the tortious aspect of the defendant's conduct, for example, failure to provide a fire escape in the hotel, or excess speed or inattentiveness while driving the car. The choice between these two approaches is a decision of policy or principle. This, however, does not make the causal inquiry itself any less factual.

Adoption of the first (overall-conduct) approach makes it easy to satisfy the actual-causation requirement, unless the defendant had no connection with any of the conditions or events that contributed to the injury. The causal inquiry therefore plays a relatively minor role in the determination of liability. Instead, liability turns primarily on policies related to deterrence of tortious behavior, allocation of losses to a tortious defendant when the losses are "within the risk" created by the tortious behavior (whether or not they were caused by the tortious aspect of the behavior), providing incentives for wealth-maximizing behavior, and so forth.¹⁰⁰

Adoption of the second (tortious-aspect) approach, however, significantly constrains pursuit of such policy goals by according a much greater role to the causal inquiry in determining tort liability. The defendant cannot be held liable unless the tortious aspect of his conduct contributed to the injury.¹⁰¹ The role of the causal inquiry becomes even more significant if, as appears to be the case, defendants are often held liable for harm caused by the tortious aspect of their behavior even though the harm was neither foreseeable nor "within the risk."¹⁰²

It is clear from the cases that the courts follow the tortious-aspect approach.¹⁰³ In fact, it is so clear that many readers may wonder why I have devoted even this many lines to belaboring the obvious. My justification is that, surprisingly, very few legal writers have noted this point,

100. E.g., G. CALABRESI, THE COSTS OF ACCIDENTS 6-7 & n.8, 131-97 (student ed. 1970); Calabresi, *supra* note 9; Green, *supra* note 8; Pound, *supra* note 7; Thode, *The Indefensible Use of the Hypothetical Case to Determine Cause in Fact*, 46 TEX. L. REV. 423 (1968).

101. For a discussion of the difficulties that this restriction creates for the wealth-maximizing theories of tort liability, see Wright, *supra* note 9.

102. See *infra* text accompanying notes 161-64.

103. E.g., Weeks v. McNulty, 101 Tenn. 495, 48 S.W. 809 (1898); Texas & Pac. Ry. v. McCleery, 418 S.W.2d 494 (Tex. 1967); Peterson v. Nielsen, 9 Utah 2d 302, 343 P.2d 731 (1959); Sowles v. Moore, 65 Vt. 322, 26 A. 629 (1893); see RESTATEMENT (SECOND) OF TORTS § 432 (1965); RESTATEMENT OF TORTS § 519 (1938).

much less emphasized its significance.¹⁰⁴ A substantial number of writers have expressly rejected the tortious-aspect approach to the cause-in-fact inquiry.¹⁰⁵ Others have written ambiguously about the need to establish a causal connection between the defendant's tortious conduct and the injury.¹⁰⁶

The proponents of the overall-conduct approach replace the tortious-aspect causation requirement with risk-related limitations on liability. One theory, championed by Leon Green, relies on ad hoc, open-ended policy considerations as part of the duty (tortious-conduct) inquiry. The other theory, developed most extensively by Robert Keeton, limits liability to harm that is "within the risks" that made the defendant's conduct tortious. This limitation occurs as part of either the tortious-conduct inquiry or the proximate-cause inquiry.

In this section I first argue that the risk theorists must smuggle the tortious-aspect causation requirement into their risk analyses to obtain

104. Those who have at least noted the point include A. BECHT & F. MILLER, *supra* note 8, *passim*; Carpenter, *Workable Rules for Determining Proximate Cause* (pt. 2), 20 CALIF. L. REV. 396, 408-19 (1932); Phillips, *Reflections on Factual Causation*, 1978 WASH. U.L.Q. 661, 663; Williams, *supra* note 21, at 64-65, 70-71, 73-75; see also H.L.A. HART & T. HONORÉ, *supra* note 8, at xxxvii-xxxviii, lviii-lxi, 117-21, 208-11, 293-94; R. KEETON, *LEGAL CAUSE IN THE LAW OF TORTS* 4-9, 12-13, 18 (1963). The economic analysts generally have assumed the tortious-aspect approach when they address the causation issue, but many have assumed the overall-conduct approach in their articles on tortious conduct. See Brown, *supra* note 59, at 327-29, 333-34 (overall conduct); Calabresi, *supra* note 9, at 79, 85-86 (tortious aspect); Calfee & Craswell, *Some Effects of Uncertainty on Compliance with Legal Standards*, 70 VA. L. REV. 965, 975, 990 & n.56 (1984) (overall conduct); Grady, *supra* note 9, at 804, 814-15, 824-25 (tortious aspect); Landes & Posner, *supra* note 9, at 111-18 (tortious aspect); Landes & Posner, *The Positive Economic Theory of Tort Law*, 15 GA. L. REV. 851, 868-70, 880-82 (1981) (overall conduct); Shavell, *supra* note 9, at 463-64, 481-82, 489 (tortious aspect); Shavell, *supra* note 59, at 10, 11 n.18, 15 (overall conduct).

105. 2 F. HARPER & F. JAMES, *supra* note 2, § 20.5, at 1138 & n.17 (But cf. *id.* § 20.2 n.4 comment, at 93-94 (Supp. 1968)); R. KEETON, *supra* note 104, at viii-ix, 10-16; Cole, *supra* note 7, at 459-60, 473-75, 482-86 & n.80, 491, 498-507 & n.133; Crowe, *The Anatomy of a Tort—Greenian, as Interpreted by Crowe, Who Has Been Influenced by Malone—A Primer*, 22 LOY. L. REV. 903, 904-05, 920-22 (1976); Epstein, *supra* note 13, at 152, 168-69, 181; Green, *supra* note 8, at 544 & n.3, 546-52, 562, 569 n.77, 576; Henderson, *A Defense of the Use of the Hypothetical Case to Resolve the Causation Issue—The Need for an Expanded, Rather than a Contracted, Analysis*, 47 TEX. L. REV. 183, 185, 195-98, 200-01, 212-13 (1969); Morris, *supra* note 3, at 1096-97, 1104; Thode, *supra* note 100, at 424-25, 428-29; see also Pedrick, *supra* note 7, at 656-57 & n.41.

Although Robert Keeton claims to be using the tortious-aspect approach, he substitutes the "harm within the risk" approach to proximate-cause limits for the tortious-aspect approach to cause-in-fact. See *infra* text accompanying notes 118-26. Keeton and Thode each report Warren Seavey's rejection of the tortious-aspect approach. R. KEETON, *supra* note 104, at 12-13; Thode, *supra* note 100, at 429 & n.23. Pedrick cites Prosser as also supporting the overall-conduct approach, but the cited portion of Prosser's text uses the defendant's negligence (excess speed) rather than his conduct as a whole (driving) as the basis for the causal inquiry. Pedrick, *supra* note 7, at 656 n.41 (citing W. PROSSER, *supra* note 1, § 41, at 237 & n.6).

106. E.g., J. FLEMING, *supra* note 4, at 170-71; C. MORRIS & C.R. MORRIS, *supra* note 4, at 154-201; W. PROSSER, *supra* note 1, § 30, at 143; *id.* § 41, at 236-43; *id.* § 42, at 249; *id.* § 45, at 289; PROSSER & KEETON, *supra* note 7, §§ 41-45, at 263-321; Smith, *Legal Cause in Actions of Tort* (pts. 1-3), 25 HARV. L. REV. 103, 223, 303 (1911-1912).

results that are consistent with the decisions of the courts. I then elaborate the tortious-aspect causation requirement and apply it to a broad range of cases to demonstrate its importance in determining liability.

1. *Green's Duty-Risk Theory*

Leon Green and his followers insist that the causal inquiry be applied to the defendant's conduct as a whole. They vigorously reject any requirement of a causal relation between the injury and the tortious aspect of the defendant's conduct. For them, the critical factors in determining liability are the relevant social policies. They claim that the court should weigh these policies in each case to decide whether the injury represents the type of risk for which the defendant should be held responsible (the duty issue). Furthermore, these policies should also be weighed by the jury, in conjunction with the other contributing factors and the foreseeability of harm of the sort that occurred, to decide whether the defendant should compensate the plaintiff (the negligence, or breach of duty, issue).¹⁰⁷

Green offers the following example to demonstrate how his approach "relieves the causation concept of most of its burdens and confusing terminology."¹⁰⁸ A driver *X* runs into *P* without trying to use his brakes. *X* was going too fast to stop anyway. Moreover, unbeknown to *X*, the brakes were defective. *D*, who knew the brakes were defective, had sold the car to *X* without disclosing the defective condition of the brakes. Green notes that, under his approach, there is a clear causal connection between each of *X*'s and *D*'s activities and *P*'s injury. "*D*'s selling the car to *X* [regardless of *D*'s knowledge of the defective brakes] contributed to *P*'s injuries. . . . *X*'s buying and driving the car [regardless of his speeding and failure to try to use the brakes] contributed to *P*'s injury."¹⁰⁹ The liability determination, therefore, turns on the scope of the duties and the question of whether either *D* or *X* was negligent with respect to *P*.¹¹⁰

Green states without elaboration that "*X* owed no duty to anyone to use the defective brakes, and did not violate any duty to anyone in failing to use them."¹¹¹ These assertions are not plausible. In light of the obvious risks, a court would almost certainly decide that *X* owed a duty to *P* and to everyone else on the road not to speed, and that *X* also owed a duty to *P* to try to use the brakes to avoid hitting *P*. A jury would

107. E.g., Crowe, *supra* note 105; Green, *supra* note 8; Thode, *supra* note 100; see also L. GREEN, W. PEDRICK, J. RAHL, E.W. THODE, C. HAWKINS, A. SMITH & J. TREECE, CASES ON THE LAW OF TORTS (2d ed. 1977) (no discussion of actual causation issue).

108. Green, *supra* note 8, at 569 n.77.

109. *Id.*

110. *Id.*

111. *Id.*

undoubtedly conclude that *X* breached his duties to *P* by speeding and failing to try to use his brakes. The implicit assumption in Green's statement is that there is no duty and no negligence if performance of the duty, in hindsight, would not have made any difference, even when the defendant was unaware that performance would not make a difference. But this is precisely the causal issue, which has been smuggled into the duty analysis.¹¹² The duty and breach of duty are clear. Under the criminal law, *X* would be liable for his speeding and failure to brake whether or not they contributed to *P*'s injury. The troublesome issue under the tort law is the causal one: did *X*'s speeding or failure to try to use his brakes (the negligent aspects of his conduct) contribute to *P*'s injury, even though the injury would have occurred anyway given the defective brakes?¹¹³ This causal issue is hardly clarified by the use of duty language and duty analysis.

The confusion introduced by trying to resolve causation problems through duty-risk analysis is further exemplified by Green's analysis of *D*'s tort liability. He notes initially that "*D* was under a duty to *P* and to *X* and other users of the highway to refrain from selling the car to *X* knowing it had defective brakes without disclosing that fact to *X*." But, he asks,

Was the risk of driving at a speed which would not permit *X* to avoid collision with *P*'s car within the scope of *D*'s duty to *X* and to *P*. . . ?

Would *X* have bought and driven the car if *D* had disclosed to him that the brakes were defective? Under the evidence this could be a jury issue. . . . [I]t cannot be presumed that he would have risked his own safety and that of other users of the highway by driving a car with defective brakes. The risk of driving the car on the highway, irrespective of the speed with which it was driven, falls within the scope of *D*'s duty to *X* and to *P* with respect to the injuries suffered by them.¹¹⁴

Green apparently concludes that *D* is liable for *any* injuries inflicted on others by *X*'s driving, no matter how negligently *X* drove, because presumably *X* would not have been driving if *D* had told him about the defective brakes. This again is a causal ("but for") argument masquerading as a duty issue. Moreover, it is a very weak causal argument. Assuming that *X* would not have bought and used the particular car if he had been told about the defective brakes, he presumably would have been driving another car: his old car, a different new car, or a leased or borrowed car. There is no reason to believe that the particular car had anything to do with *X*'s speeding or failure to brake. Thus, although *D*'s

112. For another unsuccessful attempt to replace causal analysis with tortious-conduct analysis, which also relies on causal arguments, see Landes & Posner, *supra* note 9, at 110-11, 115-16, 119-22, criticized in Wright, *supra* note 9, at 452-55.

113. The answer is yes. See *infra* text accompanying notes 274-76.

114. Green, *supra* note 8, at 569 n.77.

failure to disclose was negligent, it was not a cause of *P*'s injury.¹¹⁵

The above discussion reflects Green's final views on causation. These views, published in 1962, had evolved over several decades as he led the legal realists' fight to unburden causal analysis of all the critical policy issues. In his initial writings, he clearly used the negligent aspect of the defendant's conduct, rather than the defendant's conduct as a whole, as the starting point for the cause-in-fact inquiry.¹¹⁶ Even in his 1962 article, he continued to focus on the negligent aspect in the cases involving failure to provide fire escapes, life saving equipment, lighting for stairways, and so forth. In each of these cases, the failure to provide the safety precaution is the negligent aspect of the defendant's overall activity of operating a hotel, boat, store, train station, or whatever.¹¹⁷ These cases, and the example discussed above, demonstrate the importance of focusing the causal inquiry on the tortious aspect of the defendant's conduct. Failure to do so burdens the duty analysis with causation issues that may well be mishandled since they are not explicitly recognized as causation issues.

2. Keeton's "Harm Within the Risk" Theory

Similar problems appear in Robert Keeton's version of the risk theory of liability. Keeton employs risk analysis in the proximate-cause inquiry rather than in the duty-negligence (tortious-conduct) inquiry. Initially, he recognizes the courts' insistence that the plaintiff's injury has been caused by the negligent aspect of the defendant's conduct. The first two formulations of his Risk Rule appear to be successively more explicit statements of this requirement.¹¹⁸ However, he eviscerates the requirement by interpreting "negligent aspect" to mean the *type of risk* that

115. See *Saunders Sys. Birmingham Co. v. Adams*, 217 Ala. 621, 117 So. 72 (1928). For an analysis of this example under the causal test proposed in this Article, see *infra* text accompanying notes 274-76.

116. L. GREEN, JUDGE AND JURY 188, 229-30 (1930); L. GREEN, *supra* note 2, at 3-4, 41, 145-46, 148-49, 151-52, 164, 166, 181-84 n.60.

117. Green, *supra* note 8, at 559. *But see id.* at 560 n.45 (citing *Stacy v. Knickerbocker Ice Co.*, 84 Wis. 614, 54 N.W. 1091 (1893) (failure to provide warning fence)). Wayne Thode has applied Green's later views to the cases involving failure to provide a fire escape. He asserts that furnishing the room to the deceased is sufficient to satisfy the causal requirement, even when the failure to provide the fire escape did not contribute to the victim's death. Thode, *supra* note 100, at 433-34. Thode tries to convince those who disagree with this approach by arguing that we would surely hold the defendant liable if the deceased stayed in his room because he knew there was no fire escape, yet "*D*'s conduct has not changed." *D*'s conduct has not changed, but its causal effects have. The failure to provide the fire escape has provided the reason for the deceased to stay in his room—a causal connection that is lacking when it is assumed that the deceased stayed in his room for other reasons. See *infra* note 212.

118. The first formulation states: "A negligent actor is legally responsible for that harm, and only that harm, of which his *negligence* is a cause in fact." R. KEETON, *supra* note 104, at 4 (emphasis in original). The second formulation replaces "his *negligence*" with "the *negligent aspect of his conduct*." *Id.* at 9 (emphasis in original).

made the defendant's conduct negligent, rather than the aspects of the *conduct itself* that created that type of risk.¹¹⁹ He then restates the requirement in a third formulation that is very similar to Green's approach:

A negligent actor is legally responsible for the harm, and only the harm, that not only (1) is caused in fact by his conduct [as a whole] but also (2) is a result within the scope of the risks by reason of which the actor is found to be negligent.¹²⁰

Under this formulation of the Risk Rule, the only causal requirement is that the injury have been caused by the defendant's conduct as a whole. The second part of the formula merely requires that the injury fit the description of one of the risks that made the defendant's conduct negligent. It calls for a *comparison* of the injury with the relevant risk description (the so-called "negligent aspect" of the conduct), rather than an inquiry into the *causal connection* between the injury and some prior act, omission, or condition.¹²¹

For example, Keeton discusses a hypothetical case in which the defendant negligently placed an unlabeled can of rat poison next to food

119. Thus, in his primary example, Keeton describes the negligent aspect as "placing the poison where it was likely to be mistaken for something intended for human consumption." *Id.* at 5; *see id.* at 8-9. This is a description in terms of the risk created, rather than a description of the conduct (placing unlabeled poison near food) which created the risk. *See also id.* at 14-16, 126-27 n.11.

120. *Id.* at 10. For Keeton's distinction between the conduct as a whole and the negligent aspect of the conduct, see *id.* at 4-5.

There are some important differences between Keeton's and Green's respective approaches. For example, Keeton limits the relevant risks to those that were foreseeable at the time the defendant acted (or failed to act) and were the basis for characterizing the defendant's conduct as negligent. That is, his "harm within the risk" proximate-cause analysis builds on the foreseeability analysis employed in the determination of negligence. *Id.* at 10-11, 18-22, 51, 56. Green, on the other hand, would identify the relevant risks as part of a broad social-policy analysis of the duty issue, in which foreseeability of the risk is not necessary and is only one of many considerations. Foreseeability of the type of consequence, but not the manner of its occurrence, is an important part of Green's elaboration of the negligence (breach of duty) issue. The negligence-foreseeability formula, however, is seen merely as a useful method for presenting the ultimate issue of culpability or responsibility to the jury, which can and often does impose liability despite lack of foreseeability of the injury. Green, *supra* note 8, at 562-64, 566-68 & n.72, 570-74; Green, *Foreseeability in Negligence Law*, 61 COLUM. L. REV. 1401 (1961).

121. R. KEETON, *supra* note 104, at 48; *see id.* at 52-59. This confusion of causal analysis with "harm within the risk" analysis leads Keeton to assert that "the Risk Rule is indeed a rule of causation in a cause-in-fact sense" and that the proximate-cause inquiry is therefore a causal inquiry. *Id.* at 13; *accord*, Robinson, *supra* note 7, at 756 n.143, 757 n.144; *see* R. KEETON, *supra* note 104, at 17-18, 81. Unfortunately, this confusion now pervades the fifth edition of Prosser's hornbook, of which Keeton is a coeditor. Prosser's longstanding insistence that the actual-cause inquiry is factual and that the proximate-cause inquiry (as distinct from the preliminary issue of actual causation) is noncausal and policy-dependent has been systematically replaced by statements regarding the policy-dependent nature of the actual-cause inquiry and the causal nature of at least part of the proximate-cause inquiry. Compare W. PROSSER, *supra* note 1, § 41, at 237; *id.* § 42, at 244, 249-50; *id.* § 43, at 250-51; *id.* § 45, at 289, with PROSSER & KEETON, *supra* note 7, at xix; *id.* § 41, at 264-65; *id.* § 42, at 273, 274, 279-80; *id.* § 43, at 280-81; *id.* § 45, at 321.

on a shelf in the kitchen of his restaurant. The shelf was next to a hot stove, and the heat from the stove caused the can of poison to explode. The force of the explosion killed a delivery man. The defendant did not know, and had no reason to know, that the rat poison might explode. The defendant's conduct as a whole—placing the can on the shelf or, broader yet, operating the restaurant—was a cause of the injury. However, Keeton states, the defendant is not liable because the description of the injury ("injury by explosion") does not match the description of the risk that made the defendant's conduct negligent ("injury by poisoning").¹²²

But suppose a customer died by poisoning as a result of eating the rat poison, which was deliberately served to him by someone who knew it was poisonous. The defendant's negligence did not contribute to the injury, yet the defendant will be liable under Keeton's Risk Rule unless the risk is described more precisely, for example, "injury by *accidental* poisoning." Suppose then that a customer is poisoned accidentally and nonnegligently as a result of being fed some substance other than the rat poison. The risk must be detailed even further: "injury by accidental poisoning with the rat poison that the defendant placed on the shelf." What if a government health inspector accidentally contaminates the customer's food with the rat poison, after opening and inspecting the can, which had been properly labeled and moved away from the food several days earlier? We must add a qualifier, "as a result of mistaking the poison for food," to the previous description of the risk. Suppose the government inspector mistook the poison for food even though it was now away from food and properly labeled. We must add a further qualifier: "because the rat poison was unlabeled or near food."

At this point, we have included in the description of the risk a description of the aspect of the defendant's *conduct* which made that conduct negligent: having unlabeled poison around, or having poison (even if properly labeled) near food. Moreover, we have included in the description the requirement that the negligent aspect have been a cause of the victim's injury: "injury . . . because the rat poison was unlabeled or near food." Unless the description of the risk explicitly or implicitly includes this tortious-aspect causation requirement, the Risk Rule will encounter difficulties similar to those described in the previous paragraph. Contrary to the courts' practice, the rule will permit the defendant to be held liable even though the tortious aspect of his conduct was not a cause of the injury. The same problem exists with respect to the first and second formulations of Keeton's Risk Rule, since Keeton interprets the terms "negligence" and "negligent aspect" in them to mean the

122. R. KEETON, *supra* note 104, at 3, 10.

type of risk that made the defendant's conduct negligent, rather than the aspects of the defendant's conduct which created that type of risk.¹²³

Keeton occasionally incorporates the tortious-aspect causation requirement into the description of the risk, in order to avoid liability that otherwise would exist under the Risk Rule.¹²⁴ More often, his discussions of specific cases implicitly assume that the injury not only must have been within the broadly stated risk, but also must have been caused by the tortious aspect of the defendant's conduct or activity.¹²⁵

In sum, the Risk Rule mandates results that are inconsistent with those reached by the courts, unless the tortious-aspect causation requirement is incorporated into the description of the risk. But, as with Green's duty analysis, this awkward and circuitous method of posing the causal inquiry creates a substantial danger of misperceiving or mishandling the causal element in the liability analysis. Instead, as some proponents of the risk theory have recognized,¹²⁶ the "harm within the risk" limitation should be viewed as a (proximate-cause) supplement to the tortious-aspect causation requirement, rather than as a substitute for it.

3. *The Tortious-Aspect Causation Requirement*

As far as I am aware, only two of the many books and articles on causation in tort law have emphasized the existence and significance of

123. See *supra* note 119.

124. For example, Keeton discusses the case in which the victim negligently sat on an unsafe (weak) wall, but was injured when the defendant crashed into the wall with a force that would have knocked the wall down even if it had been safe. He describes the victim's negligence, in risk language, as "placing himself where he was likely to be injured by the collapse of the wall, either without an external impact or under an external impact insufficient to cause the collapse of a safe wall," and he notes that "[t]his aspect of his conduct was not a *sine qua non* of the injury he suffered." R. KEETON, *supra* note 104, at 127 n.11.

Of course, the victim's placing himself on the wall was a *sine qua non* of the injury. What Keeton means is that the risk which made his conduct negligent did not materialize. To reach this result, Keeton must define the risk more specifically than "injury due to collapse of the wall." His definition must include the requirement that the collapse be caused by the unsafe condition of the wall: the tortious-aspect causation requirement. (The description of the tortious aspect of the defendant's conduct must include the unsafe condition of the wall, but the unsafe condition was not a *sine qua non* of the injury. Nevertheless, it may have contributed to the injury. See *infra* text accompanying notes 268-73.)

125. See R. KEETON, *supra* note 104, at 6, 10, 49-78.

126. E.g., Williams, *supra* note 21, at 62-65, 70-71, 73-75, 79; Williams, *The Risk Principle*, 77 LAW Q. REV. 179, 179 n.4 (1961); see also Malone, *supra* note 27, at 370-74. For a spirited exchange between a proximate-cause risk theorist and a duty-negligence risk theorist, see Henderson, *supra* note 105, and Thode, *A Reply to the Defense of the Use of the Hypothetical Case to Resolve the Causation Issue*, 47 TEX. L. REV. 1344 (1969). While both Henderson and Thode insist that the cause-in-fact inquiry is concerned only with the causal effect of the actor's conduct as a whole, they each correctly accuse the other of relying heavily on the tortious-aspect causation requirement in his risk analysis, and they note the unnatural and awkward language that the other is required to use when the causal inquiry is incorporated into the risk analysis.

the tortious-aspect causation requirement, and in each instance the tortious-aspect analysis is applied only to negligence cases.

The first of these is Charles Carpenter's extremely illuminating article,¹²⁷ in which he attempts, "without imposing a preconceived theory, to restate, in consistent and workable form the principles [of actual cause and proximate cause] . . . actually found in the decisions."¹²⁸ Although in a few places he confuses the "harm within the risk" issue with the negligent-aspect causation requirement,¹²⁹ he subsequently clearly distinguishes the two issues and demonstrates the pervasiveness and greater significance of the negligent-aspect causation requirement.¹³⁰ However, he does not explain how one identifies the negligent aspect of conduct.

The second is Arno Becht and Frank Miller's book on actual-causation analysis in tort law. Becht and Miller suggest that the negligent aspect of a person's conduct should be identified as the minimum aspect which would need to be changed to make the conduct nonnegligent.¹³¹ The usefulness of this approach in clarifying the liability analysis is demonstrated throughout their book.¹³² However, their conception of the negligent aspect of conduct is too broad. They argue that the negligent aspect is always an unqualified act or omission. They do not further specify the negligent aspect by including the circumstances or conditions which have to be combined with the act or omission in order to characterize the conduct as negligent.

For example, Becht and Miller discuss a case in which the plaintiff negligently sat on an unsafe (weak) wall, but was injured when the defendant knocked down the wall with his car. They insist that the negligent aspect of the plaintiff's conduct was the act of sitting on the wall.

127. Carpenter, *Workable Rules for Determining Proximate Cause* (pts. 1-3), 20 CALIF. L. REV. 229, 396, 471 (1932).

128. *Id.* at 257-58.

129. *Id.* at 231 (loaded gun discussion); *id.* at 231 n.5 (*Butz* case); *id.* at 253-54 (*Teis* case).

130. *Id.* at 408-19, 471-539.

131. A. BECHT & F. MILLER, *supra* note 8, at 34; *see also id.* at 12-13, 27-28 & n.40, 87-90.

132. *E.g., id.* at 34-42, 54, 58, 59-61, 85, 90-91, 140, 169-86. The authors' discussion of *Kernan v. American Dredging Co.*, 355 U.S. 426 (1958), is particularly interesting. The defendant's act of carrying a kerosene signal lantern at a height lower than that mandated by statute was deemed sufficient to hold him liable for the ignition of petroleum vapors on the surface of the water, even though the purpose of the statute was to prevent collisions between ships, not the ignition of surface vapors. Becht and Miller, expanding upon a footnote by Justice Harlan, 355 U.S. at 442 n.1, note that the negligence was the omission of a signal lantern at the required height, not the presence of one at the lower height (which was not in itself prohibited). If there had been a lantern at the required height in addition to the one at the lower height, there would have been no negligence (breach of the statute), yet the vapors would still have been ignited. Thus, the negligent aspect was not a cause of the injury. A. BECHT & F. MILLER, *supra* note 8, at 38-41. *But see H.L.A. HART & T. HONORÉ, supra* note 8, at 289 (failing to note this point). If the defendant's attorney had correctly identified the negligent aspect and stressed the lack of causal connection between the negligence and the injury, rather than conceding negligent causation and relying on the "harm outside the risk" argument, he probably would have won the case.

This clearly was a necessary condition for the plaintiff's injury. They therefore conclude that the negligent-aspect causation requirement is satisfied. They refuse to consider, as part of the causal inquiry, whether the unsafe condition of the wall contributed to the injury, since the condition of the wall was not a part of the plaintiff's conduct.¹³³

But the unsafe condition is a necessary element in the description of the plaintiff's *negligent* conduct. The act of "sitting on a wall" is not in itself negligent. The act of "sitting on an unsafe (weak, structurally unsound) wall" is negligent, when it is coupled with the mental element required for negligence (actual or constructive knowledge of the unsafe condition). The act of sitting and the unsafe condition are *both* necessary to make the plaintiff's conduct negligent. Therefore it must be established that both the act and the condition contributed to the injury before it can be said that the plaintiff's negligence contributed to the injury.¹³⁴

In general, the tortious aspect of a person's conduct or activity is a cause of an injury only if each of its necessary elements (act, omission, condition, or circumstance) contributed to the occurrence of the injury.¹³⁵ If a certain element did not contribute to the injury, but was necessary to make the conduct or activity tortious, then it cannot be said that the tortious aspect of the conduct or activity was a cause of the injury.¹³⁶

The actor's mental state (intent, or actual or constructive knowledge of a certain risk), which is required in addition to the conduct or activity itself in order for the conduct or activity to be considered tortious, does not enter directly into the causal inquiry. Instead, the intent or knowledge is used to determine which acts, omissions, conditions, and circumstances constitute the tortious aspect of the conduct or activity. For example, in the explosive rat poison case, the actor was aware (or should have been aware) of its poisonous nature, but was not aware (and had no reason to be aware) of its explosive nature. Therefore, the poisonous nature of the substance was a constituent of the tortious aspect of the actor's conduct, but its explosive nature was not.

Becht and Miller's overly broad conception of the tortious aspect of conduct forces them erroneously to concede negligent causation and to

133. A. BECHT & F. MILLER, *supra* note 8, at 183-84; see also J. FLEMING, *supra* note 4, at 192 & n.43; Malone, *supra* note 27, at 371.

134. H.L.A. HART & T. HONORÉ, *supra* note 8, at 210; see *infra* text accompanying notes 268-73.

135. When there is more than one tortious aspect, each must be considered, and the tortious-aspect causation requirement is satisfied if any of them contributed. For example, in the rat poison case, there were two tortious aspects: (1) having unlabeled poison around and (2) having poison (even if labeled) near food. If either tortious aspect contributed to the injury, the tortious-aspect causation requirement is satisfied.

136. See H.L.A. HART & T. HONORÉ, *supra* note 8, at 117-20, 208-11; Carpenter, *supra* note 127, at 409-10.

rely instead on the "harm outside the risk" rationale to limit liability in a number of cases.¹³⁷ It also leads them to adopt Green's and Keeton's approach to causation of injury in the strict liability cases. They assert that, under strict liability, the causal inquiry is applied to the defendant's conduct as a whole.¹³⁸ Presumably, they again would use the "harm outside the risk" argument to limit liability.

Properly understood, the tortious-aspect causation requirement applies to all tort actions, not just to negligence actions.

In the traditional strict liability cases, the tortious aspect of an activity includes all the acts, omissions, and conditions necessary to make it ultrahazardous or abnormally dangerous: for example, the keeping or use of a highly explosive substance, a poisonous fugitive substance, a highly radioactive or toxic substance, a vicious or wild animal, a foraging animal, or a large bulk accumulation of liquid. That tortious aspect must be a cause of the injury, as the original *Restatement of Torts* made clear: "[O]ne who carries on an ultrahazardous activity is liable to another . . . for harm resulting thereto from that which makes the activity ultrahazardous."¹³⁹

Unfortunately, however, the academic proponents of the risk theory succeeded in rewording section 519 in the *Restatement (Second)* to make it conform to Keeton's Risk Rule: "(1) One who carries on an abnormally dangerous activity is subject to liability for harm to [another] . . . resulting from the activity . . . (2) This strict liability is limited to the kind of harm, the possibility of which makes the activity abnormally dangerous."¹⁴⁰ This Risk Rule formula will produce unintended results

137. A. BECHT & F. MILLER, *supra* note 8, at 61-62 (plaintiff negligently stood on unrailed portion of icy platform, but was injured by wall falling on that portion of platform rather than by slipping on ice); *id.* at 150-51, 199 (accident occurs while plaintiff or defendant is engaged in secular activity in violation of Sunday blue law). See also their confusing discussion of the license cases at *id.* at 141-50, 194-99.

138. *Id.* at 46, 48, 168.

139. RESTATEMENT OF TORTS § 519 (1938) (emphasis added).

140. RESTATEMENT (SECOND) OF TORTS § 519 (1977) (emphasis added). Among the Advisers to the Reporter for the *Restatement (Second)* were such prominent academic proponents of the risk theory as Eldredge, James, Robert Keeton, Page Keeton, Malone, and Seavey.

The *Restatement (Second)* also dropped the language in the original *Restatement* which limited liability to those whom the actor "should recognize as likely to be harmed" by the miscarriage of the activity. However, the drafters apparently assumed that this limitation was retained by the "kind of harm" language in § 519(2). Thus, comment e to § 519(2) refers to risk and injury "to those in the vicinity" of a possible explosion, and the accompanying illustration 1 rejects liability when the actor "has no reason to know of the presence of B's mink ranch nearby," even though "[t]he noise of the blasting frightens the mink and the fright causes them to kill their young."

The mink illustration is based on *Madsen v. East Jordan Irrigation Co.*, 101 Utah 552, 125 P.2d 794 (1942). The result seems inconsistent with RESTATEMENT (SECOND) OF TORTS § 522(b) (liability despite unforeseeable contributing action of animal) and also the Risk Rule itself. Frightened or startled reactions to the noise of blasting, as well as direct effects of the force of the blast, are foreseeable and "within the risk," and the peculiar reaction of the mink in such situations

unless the tortious-aspect causation requirement is read back into it or into the definition of each particular risk ("kind of harm").¹⁴¹ For example, suppose a truck transporting dynamite nonnegligently runs over a detonation device set up by someone else, thereby causing explosives belonging to that other person, but not those in the truck, to go off and cause injury in the vicinity of the truck.¹⁴² Unless the tortious-aspect causation requirement is read in, the owner of the truck will be strictly liable under section 519. Once again, the "harm within the risk" principle can be at best a supplement to, rather than a substitute for, the tortious-aspect causation requirement.

In strict product liability cases, the tortious aspect of the defendant's conduct or activity is the defect in the particular product, and it is clear that the defect must be a cause of the injury.¹⁴³

In the intentional tort cases, the tortious aspect is all the acts, omissions, conditions, and circumstances necessary to create the substantial certainty of invasion of a protected interest, or the acts or omissions done for the purpose of causing such an invasion. The substantial-certainty variety of intentional tort is analogous to negligence, with one difference. The risk created in the intentional tort action must be substantially certain to be realized but need not be unreasonable, while the risk created in the negligence action can be much less likely but must be unreasonable.¹⁴⁴

Finally, in the nuisance cases, liability is based on conduct that is either negligent, abnormally dangerous, or intentional (usually in the sense of substantially certain although reasonable).¹⁴⁵ Therefore, the tor-

is a well-established characteristic. See, e.g., *Wildwood Mink Ranch v. United States*, 218 F. Supp. 67 (D. Minn. 1963); *Gronn v. Rogers Constr.*, 221 Or. 226, 350 P.2d 1086 (1960); *Summit View, Inc. v. W.W. Clyde & Co.*, 17 Utah 2d 26, 403 P.2d 919 (1965); *Foster v. Preston Mill Co.*, 44 Wash. 2d 440, 268 P.2d 645 (1954); *MacGibbon v. Robinson*, [1952] 4 D.L.R. 142 (B.C.); R. KEETON, *supra* note 104, at 135 n.100. In the *Madsen* case, the defendant irrigation company presumably was familiar with the area and, at any rate, was blasting only 100 yards from the mink ranch. Both the vibrations and the noise terrified the mother mink, which reacted in a natural way (for mink). The result is justified only as an application of the principle that a person need not take excessive care to avoid injury to extrasensitive plaintiffs and, correspondingly, should not be held strictly liable for "harm [that] would not have resulted but for the abnormally sensitive character of the plaintiff's activity." RESTATEMENT (SECOND) OF TORTS § 524A (1977).

141. See *supra* text accompanying notes 122-23.

142. Cf. RESTATEMENT (SECOND) OF TORTS § 519 comment e (1977); R. KEETON, *supra* note 104, at 105.

143. *Greenman v. Yuba Power Prods.*, 59 Cal. 2d 57, 62-64, 377 P.2d 897, 900-01, 27 Cal. Rptr. 697, 700-01 (1962); *Codling v. Paglia*, 32 N.Y.2d 330, 342, 298 N.E.2d 622, 628-29, 345 N.Y.S.2d 461, 469-70 (1973); RESTATEMENT (SECOND) OF TORTS § 402A (1965).

144. Thus, it is not true, as Keeton claims, that "[n]o concept of risk is resorted to in determining the issue of liability for intentional tort." R. KEETON, *supra* note 104, at 100-01.

145. *Boomer v. Atlantic Cement Co.*, 26 N.Y.2d 219, 257 N.E.2d 870, 309 N.Y.S.2d 312 (1970); *Morgan v. High Penn Oil Co.*, 238 N.C. 185, 77 S.E.2d 682 (1953); *Jost v. Dairyland Power*

tious aspect is defined according to the type of tortious conduct involved in each case.

4. *The Power of the Tortious-Aspect Causation Requirement*

The tortious-aspect causation requirement resolves many of the cases that have been prominent in discussions of the proximate-cause limits on tort liability, especially in the writings of the risk theorists. There is thus no need to resort to noncausal policies or principles under the rubrics of proximate cause, duty, recoverable damages, and so forth in order to deny liability in these cases. In each of the following examples, the italicized condition was a necessary constituent of the tortious aspect of the defendant's or the plaintiff's conduct, but the condition did not contribute to the injury, so the tortious aspect was not a cause of the injury. The defendant handed a *loaded* gun to a child, who dropped the gun on his foot.¹⁴⁶ The defendant's box of *explosives* fell on the plaintiff's foot, or was being transported in a truck that ran into the plaintiff (no explosion occurred).¹⁴⁷ The plaintiff stood on the *icy, unrailed* portion of a platform, where a wall fell on him.¹⁴⁸ The plaintiff stood in an area *where trucks sped through*, but was hit by a railroad freight car that came through the wall.¹⁴⁹ The defendant tested an unguarded fan on the floor where men *with loose clothing* had to bend over it (creating a risk that the clothing might get caught in the fan), but the plaintiff stuck his fingers into the fan.¹⁵⁰

Coop., 45 Wis. 2d 164, 172 N.W.2d 647 (1969); RESTATEMENT (SECOND) OF TORTS §§ 822, 825, 826 (1977); *see also* Renken v. Harvey Aluminum, 226 F. Supp. 169 (D. Or. 1963).

146. RESTATEMENT (SECOND) OF TORTS § 281 illustration 3 (1965); L. ELDREDGE, MODERN TORT PROBLEMS 18 (1941); J. FLEMING, *supra* note 4, at 192; 2 F. HARPER & F. JAMES, *supra* note 2, § 20.5, at 1136-37; Carpenter, *supra* note 127, at 231.

147. RESTATEMENT (SECOND) OF TORTS § 519 comment e (1977); Seavey, *Mr. Justice Cardozo and the Law of Torts*, 39 COLUM. L. REV. 20, 35, 52 HARV. L. REV. 372, 387, 48 YALE L.J. 390, 405 (1939); Williams, *supra* note 21, at 70-71.

148. Smithwick v. Hall & Upson Co., 59 Conn. 261, 21 A. 924 (1890), *discussed in* A. BECHT & F. MILLER, *supra* note 8, at 61-63; L. ELDREDGE, *supra* note 146, at 21-22; H.L.A. HART & T. HONORÉ, *supra* note 8, at 209.

149. Gray v. Scott, 66 Pa. 345 (1870), *discussed in* L. ELDREDGE, *supra* note 146, at 22. Almost all the cases discussed by Eldredge in his argument on behalf of the risk theory involve a lack of causation by the tortious aspect of the actor's conduct. In addition to the *Gray* case, see the examples cited *supra* notes 146 & 148 and Falk v. Finkelman, 268 Mass. 524, 168 N.E. 89 (1929) (car parked *overtime* was knocked into another car by fire engine); New York, L.E. & W.R. Co. v. Ball, 53 N.J.L. 283, 21 A. 1052 (1891) (passenger in *baggage* compartment in violation of rule was injured in train wreck; injury unaffected by being near baggage); Hudson v. Lehigh Valley R.R., 54 Pa. Super. 107 (1913) (bicyclist who *failed to "stop, look and listen"* for train was hit by gate negligently lowered onto him). These cases are all discussed in a section entitled "The Hazard Problem," in L. ELDREDGE, *supra*, at 17-24, which also is the title of RESTATEMENT (SECOND) OF TORTS § 281 comment e, for which Eldredge was an Adviser. *See supra* note 140.

150. Thurogood v. Van den Berghs & Jurgens Ltd., [1951] 2 K.B. 537 (C.A.), *discussed in* Goodhart, *The Imaginary Necktie and the Rule in Re Polemis*, 68 LAW. Q. REV. 514 (1952); Williams, *supra* note 21, at 71-72. Goodhart and Williams assume that the loose-clothing element

The tortious-aspect causation requirement similarly resolves many of the troublesome statutory-violation (negligence per se) cases. Initially, it should be noted that many statutes are not passed for safety purposes¹⁵¹ or are not meant to lay down rigid rules of behavior, but rather are meant to be administered flexibly. In such cases, there is no basis for inferring negligence from violation of the statute, so the issues of actual cause and proximate cause never arise. Neither the tortious-aspect causation requirement nor the "harm within the risk" limitation is relevant or needed to deny liability. However, even if the violation is erroneously treated as tortious conduct, the tortious-aspect causation requirement rarely will be satisfied. For example, if an injury occurred while the defendant or the plaintiff was engaged in an activity *on Sunday* (in violation of a Sunday blue law), had parked his car at a meter *without paying*, or was hunting *without a permit*, the italicized element in each instance was a necessary constituent of the statutory violation but did not contribute to the injury. Therefore, the "tortious" aspect of the conduct was not a cause of the injury.¹⁵² The few cases which nevertheless impose liability on the defendant or refuse to allow recovery by the plaintiff because of the statutory violation ignore tort liability principles in order to further the purposes of the criminal law. These cases, however, generally have been repudiated.¹⁵³

The failure to satisfy the tortious-aspect causation requirement also explains the denial of liability in many cases where the statute was passed for safety purposes, but a necessary element of the illegal or tortious aspect of the defendant's conduct did not contribute to the injury. For example, Prosser cites cases in which the defendant set out a *poisonous* substance, which exploded; or failed to erect a fence *next to a railroad track*, and a cow died from eating too much; or sold *explosive* materials (fireworks), which a child ate.¹⁵⁴

was necessary to the finding of negligence, but this is not clear. If "unguarded operating fan on floor near workmen" was sufficient for negligence, the tortious-aspect causation requirement was satisfied and liability seems appropriate, absent contributory negligence. See H.L.A. HART & T. HONORÉ, *supra* note 8, at 118-19.

151. W. PROSSER, *supra* note 1, § 36, at 192-93, 196.

152. H.L.A. HART & T. HONORÉ, *supra* note 8, at 210-11, 293-94; Carpenter, *supra* note 127, at 409-10, 414-15. Writers who assert that the illegal conduct was a cause of the injury erroneously apply the causal inquiry to the conduct as a whole rather than to the illegal aspect of such conduct. E.g., A. BECHT & F. MILLER, *supra* note 8, at 150-51, 199; L. GREEN, JUDGE AND JURY 234-35 (1930); W. PROSSER, *supra* note 1, § 36, at 193.

153. 2 F. HARPER & F. JAMES, *supra* note 2, § 17.6, at 995-97; H.L.A. HART & T. HONORÉ, *supra* note 8, at 210-11; W. PROSSER, *supra* note 1, § 36, at 193 n.53, 202-03. But see G. CALABRESI, *supra* note 100, at 267-74, 301-08; Cole, *supra* note 7, at 459-62, 598-607; Green, *supra* note 8, at 569 n.77; Morris, *supra* note 3, at 1096-97; Thode, *supra* note 126, at 1348-49, 1351-52; Thode, *supra* note 100, at 430-31. For a cogent critique of Thode's argument, see Strachan, *supra* note 7, at 389-91.

154. W. PROSSER, *supra* note 1, § 36, at 195-96 & n.77.

In the above cases, it can also be said that the harm was not "within the risk." The "harm within the risk" theory, however, cannot account for some other cases. Thus, if an unlicensed driver causes an automobile accident or an unlicensed medical practitioner causes a medical mishap, the harm is the sort which the legislature meant to prevent through its licensing requirements. To describe the risk as restricted to "injury at the hands of incompetents"¹⁵⁵ is to rely on and incorporate the tortious-aspect causation requirement. The tortious aspect is "driving or practicing medicine *without a license*." Although the overall conduct of driving or practicing medicine contributed to the injury,¹⁵⁶ the failure to have the required piece of paper (the license) did not.¹⁵⁷

A similar problem exists when the defendant is violating the statutory speed limit, but a child darts in front of his car whom he could not avoid even if he were driving at the legal speed. The injury in such cases, as even Green admits,¹⁵⁸ is "within the risks" meant to be avoided by the speed limit. Some argue that the injury is not within the risks, which are said to be limited to those involving "lack of control."¹⁵⁹ But the risks meant to be avoided include not only literal loss of physical control of the car but also inability to swerve or stop in time to avoid hitting something or someone. The injury is "within" this latter risk, but there is no liability since the tortious aspect (excess speed) did not contribute to the injury. Liability can be avoided under the "harm within the risk" approach only by incorporating the tortious-aspect causation requirement in the description of the risk.¹⁶⁰

When the tortious-aspect causation requirement is satisfied in a particular case, the courts often hold the defendant liable even though the injury was unforeseeable or "outside the risk," especially when the tor-

155. *Id.* at 196; PROSSER & KEETON, *supra* note 7, § 36, at 226; see Green, *supra* note 8, at 547-48.

156. PROSSER & KEETON, *supra* note 7, § 36, at 223-24; Williams, *supra* note 21, at 74-75; see A. BECHT & F. MILLER, *supra* note 8, at 141-42.

157. H.L.A. HART & T. HONORÉ, *supra* note 8, at 117-20; Carpenter, *supra* note 127, at 412-13, 414. A legislature or court might decide, however, that the lack of license should be treated as some evidence or even as *prima facie* evidence that the driver or practitioner was incompetent to engage in the activity and that the injury was caused by some unspecified instance of that incompetence (negligence). *E.g.*, N.Y. CIV. PRAC. LAW § 4504(d) (McKinney Supp. 1984); 2 F. HARPER & F. JAMES, *supra* note 2, § 20.3, at 1120-21; Williams, *supra* note 21, at 74-75. *But see* A. BECHT & F. MILLER, *supra* note 8, at 143-50.

158. Green, *Are There Dependable Rules of Causation?* 77 U. PA. L. REV. 601, 619 (1929).

159. *E.g.*, 2 F. HARPER & F. JAMES, *supra* note 2, § 20.5, at 1138 & n.16, 1148; *id.* § 20.6, at 1156 n.23.

160. See Cole, *supra* note 7, at 766-67; Henderson, *supra* note 105, at 189-92, 195-97; Thode, *supra* note 126, at 1345-47; Thode, *supra* note 100, at 427-28, 430-31. To avoid the argument that the excess speed did contribute to the injury by bringing the defendant to the precise place where the child darted out at the precise time that he darted out, I assume here that the defendant had just begun speeding. Otherwise, the tortious-aspect causation requirement is satisfied, and the defendant must rely on noncausal proximate-cause arguments to avoid liability.

tious conduct was intentional or negligent.¹⁶¹ Even the risk-theory proponents acknowledge this.¹⁶² In the statutory context, this approach is reflected in expansive interpretations of the statutory purposes,¹⁶³ or occasionally even by frank imposition of liability for injuries conceded to be "outside the risk."¹⁶⁴ Thus, the causal inquiry, focused on the tortious aspect of the defendant's conduct, plays an extremely significant role in both establishing and limiting legal responsibility.

II

THE NATURE AND CONTENT OF THE CAUSAL INQUIRY

In Part I of this Article, I analyzed the role that the causal inquiry occupies in tort-liability analysis. In this Part I examine the nature and content of the causal inquiry itself. As will be seen, efforts to articulate a comprehensive, factual test of actual causation have failed repeatedly. The causal inquiry therefore increasingly has been viewed by academics as just one more manipulable policy tool that has little or no inherent substantive content.¹⁶⁵

In this Part I attempt to rebut this prevailing view by elaborating a causal test that incorporates the traditional Humean philosophic account of the meaning of causation, as modified by John Stuart Mill. This test, which I call the "NESS" (Necessary Element of a Sufficient Set) test, was first suggested by Hart and Honoré. As elaborated here, the NESS test states that a particular condition was a cause of a specific consequence if and only if it was a necessary element of a set of antecedent actual conditions that was sufficient for the occurrence of the consequence. A substantial portion of this Part is devoted to developing this test and applying it to resolve the problematic causation cases that have resisted solution under all the previously proposed tests.

First, however, the previous tests are examined in order to provide a basis for comparison with the NESS test, to introduce the various types of problematic causation cases, and to develop some of the concepts and

161. *E.g.*, *In re Kinsman Transit Co.*, 338 F.2d 708, 723-25 (2d Cir. 1964), *cert. denied*, 380 U.S. 944 (1965); H.L.A. HART & T. HONORÉ, *supra* note 8, at 176-78, 254-90; W. PROSSER, *supra* note 1, §§ 43-44, at 250-80; Smith, *supra* note 106, at 127-28, 233-52, 321-27. Some proximate-cause limitations on liability do exist. The primary instances are when the injury would not have occurred but for unforeseeable tortious conduct by others or independent abnormal events or conditions. See H.L.A. HART & T. HONORÉ, *supra* note 8, *passim*; Carpenter, *supra* note 127, at 471-539.

162. *See, e.g.*, L. GREEN, *supra* note 2, at 177-85; 2 F. HARPER & F. JAMES, *supra* note 2, § 20.5, at 1139-51; *id.* § 20.6, at 1155-56, 1160-61; R. KEETON, *supra* note 104, at 28-32, 39, 49-54, 60-78, 96-97, 100-03, 109-10, 117; Green, *Foreseeability in Negligence Law*, 61 COLUM. L. REV. 1401, 1417-24 (1961); Williams, *The Risk Principle*, 77 LAW Q. REV. 179, 181-87, 193-203 (1961).

163. 2 F. HARPER & F. JAMES, *supra* note 2, § 17.6, at 1004-05; PROSSER & KEETON, *supra* note 7, § 36, at 226-27.

164. *See* the discussion of the *Kernan* case *supra* note 132.

165. *See* sources cited *supra* notes 6, 7 & 9.

arguments that underlie the NESS test. The first section discusses the traditional "but for" (necessary condition) test and distinguishes the two types of overdetermined-causation cases that the test fails to handle properly: the duplicative-causation cases and the preemptive-causation cases. Subsequent sections discuss modifications of or alternatives to the but-for test.

Finally, in the last section, the steps involved in applying the NESS test are analyzed in detail to demonstrate that, contrary to the currently popular view, the hypothetical inquiry associated with the but-for necessary-condition analysis does not undermine the factual nature of the causal inquiry. The miscellaneous arguments advanced by Malone to establish the policy-dependent nature of the causal inquiry are also examined and rejected.

A. *The But-For Test and Its Limits: Duplicative and Preemptive Causation*

The most widely used test of actual causation in tort adjudication is the but-for test, which states that an act (omission, condition, etc.) was a cause of an injury if and only if, but for the act, the injury would not have occurred. That is, the act must have been a necessary condition for the occurrence of the injury. The test reflects a deeply rooted belief that a condition cannot be a cause of some event unless it is, in some sense, necessary for the occurrence of the event. This view is shared by lawyers, philosophers, scientists, and the general public.¹⁶⁶

In the vast majority of cases, the but-for test works quite well as a test of actual causation. But in certain types of cases, it results in a finding of no causation even though it is clear that the act in question contributed to the injury. These are cases of *overdetermined causation*: cases in which a factor other than the specified act would have been sufficient to produce the injury in the absence of the specified act, but its effects either (1) were preempted by the more immediately operative effects of the specified act or (2) combined with or duplicated those of the specified act to jointly produce the injury.

I will refer to the first type of situation as a case of *preemptive causation*. For example, *D* shoots and kills *P* just as *P* was about to drink a cup of tea that was poisoned by *C*. *D*'s shot was a preemptive cause of *P*'s death; *C*'s poisoning of the tea was not a cause because its potential effects were preempted. I will refer to the second type of situation as a case of *duplicative causation*. For example, *C* and *D* independently start separate fires, each of which would have been sufficient to destroy *P*'s

166. E.g., T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 94-96, 98, 107-17, 119, 131-37, 139-45, 297; H.L.A. HART & T. HONORÉ, *supra* note 8, at 15-16, 69, 110; W. PROSSER, *supra* note 1, at 237-39; Mackie, *supra* note 21, at 16-21, 25-27; Williams, *supra* note 21, at 63-64.

house. The fires converge and together burn down the house. Each fire was a duplicative cause of the destruction of the house. Yet, application of the but-for test would result in a finding that *D*'s shot was not a cause of *P*'s death in the first example and that neither *C*'s nor *D*'s fire was a cause of the destruction of *P*'s house in the second example.

A different sort of objection to the but-for test focuses on the hypothetical or counterfactual nature of the inquiry that the test calls for. A number of writers have asserted that the hypothetical nature of the but-for inquiry necessarily involves or at least invites introduction of policy considerations into a supposedly factual inquiry.¹⁶⁷

While it might be thought that these difficulties could be avoided by substituting a sufficient condition test for the but-for necessary-condition test, this substitution would eliminate almost every potential cause, since few if any acts are sufficient by themselves to produce any particular consequence. Thus, in the scientific and philosophic literature on causation, it is usually stressed that the cause of an event must include *all* the conditions which together are sufficient to produce the consequence.¹⁶⁸ Consequently, both lawyers and philosophers often conclude that this scientific or philosophic concept of causation is of little relevance or use in the law.¹⁶⁹

On the other hand, the sufficient-condition test could be interpreted to mean any condition that is sufficient in combination with other conditions to produce the consequence, even though it is not sufficient by itself. Under this interpretation, however, anything could be treated as a cause simply by adding it to an already sufficient set of conditions.¹⁷⁰

Judges and legal writers have responded to the actual and perceived deficiencies in the but-for test in several different ways. A few writers have tried to improve the but-for test by modifying the manner in which it is applied.¹⁷¹ Most judges and writers, however, have adopted the substantial-factor formula, either as a supplement to or as a substitute for the but-for test depending on their degree of dissatisfaction with the but-for test.¹⁷² Others have relied on an undefined and irreducible notion of

167. E.g., J. FLEMING, *supra* note 4, at 172; Cole, *supra* note 7; Malone, *supra* note 6, at 67-68; see also sources cited *infra* note 284.

168. E.g., T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 85, 291; H.L.A. HART & T. HONORÉ, *supra* note 8, at 17-18, 21 & n.16; W. PROSSER, *supra* note 1, § 41, at 237 n.8; Williams, *supra* note 21, at 63, 66.

169. E.g., RESTATEMENT (SECOND) OF TORTS § 431 comment a (1965); T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 283-94; A. BECHT & F. MILLER, *supra* note 8, at 9, 12; J. FLEMING, *supra* note 4, at 171 & n.2, 192 & n.43; H.L.A. HART & T. HONORÉ, *supra* note 8, at 2, 9-16, 68-69; Calabresi, *supra* note 9, at 69-70, 105-06; Epstein, *supra* note 13, at 160-61; Morris, *supra* note 3, at 1088; Williams, *supra* note 21, at 66-68, 75-76.

170. Sosa, *Introduction*, in CAUSATION AND CONDITIONALS 1-3 (E. Sosa ed. 1975).

171. See *infra* text accompanying notes 176-89.

172. See *infra* text accompanying notes 190-208.

directly observable causal contribution.¹⁷³

Each of these approaches will be discussed in subsequent sections of this Part. It will be demonstrated that, under each approach, the analysis ultimately must fall back on the but-for test in order to be useful and coherent. The necessary-condition element, in some form, indeed seems to be fundamental to the concept of causation. But it obviously is too restrictive when applied in the usual way to the overdetermined-causation cases. Some writers treat the overdetermined-causation cases as genuine instances of causation that must be admitted as unexplained exceptions to the but-for test.¹⁷⁴ A larger and growing number of writers treat the overdetermined-causation cases as policy-based exceptions to the actual-causation requirement, which is equated with the but-for test.¹⁷⁵

B. Efforts to Modify the But-For Test

As mentioned above, the but-for test works well as a test of actual causation in the vast majority of cases, but produces obviously incorrect results in the overdetermined-causation cases. A few writers have tried to extend the but-for test's range of application by modifying the way in which it is applied. Rollin Perkins, Arno Becht, and Frank Miller apply the test to a detailed description of the injury or the manner of its occurrence. Glanville Williams excludes consideration of hypothetical facts. And the editors of the fifth edition of Prosser's hornbook on tort law apply the test to multiple potential causes in the aggregate. None of these approaches works. Several of them assume the very answer the test is supposed to provide.

1. Detailing the Manner of Occurrence

Perkins tries to resolve the causal issue in the overdetermined-causation cases by detailing the manner in which the injury occurred: "Whenever *that* would not have happened *when and as* it did happen, had it not

173. See *infra* text accompanying notes 209-25.

174. E.g., 2 F. HARPER & F. JAMES, *supra* note 2, § 20.2, at 1110; *id.* § 20.3, at 1122-23; W. PROSSER, *supra* note 1, § 41, at 239-40; see also Carpenter, *supra* note 127, at 397-98, 402-03, 406-07.

175. E.g., C. MORRIS & C.R. MORRIS, *supra* note 4, at 187-90; PROSSER & KEETON, *supra* note 7, § 41, at 265 & n.8, 268-69; Calabresi, *supra* note 9, at 86-87; Delgado, *supra* note 7, at 886-87, 891-92; Edgerton, *supra* note 3, at 373; Grady, *supra* note 9, at 804 & n.21; Malone, *supra* note 6, at 89-97; McLaughlin, *Proximate Cause*, 39 HARV. L. REV. 149, 153-54 (1925); Peaslee, *Multiple Causation and Damage*, 47 HARV. L. REV. 1127, 1130-32 (1934); Rosenberg, *supra* note 7, at 855 & n.27, 863-64; Strachan, *supra* note 7, at 389-95; Williams, *supra* note 21, at 75-76. Some writers, who rely on tests other than the but-for test, treat only the most difficult overdetermined-causation cases as policy-based exceptions to the actual-causation requirement. A. BECHT & F. MILLER, *supra* note 8, at 95-98, 210-11; J. FLEMING, *supra* note 4, at 174 n.19; H.L.A. HART & T. HONORÉ, *supra* note 8, at 7-8, 125, 239-41; Pedrick, *supra* note 7, at 645, 654, 658-59; Weinrib, *supra* note 7, at 529-33.

been for *this*, *this* is an actual cause of *that*.¹⁷⁶ For example, if a victim is killed instantaneously by a bullet wound while already dying from a knife wound, he would not have died, as he in fact did, at the earlier moment and by a bullet wound if not for the firing of the bullet. Without the bullet wound, he would have died later, from the effects of the knife wound. This, however, did not in fact happen. Thus, the bullet wound, but not the knife wound, is a but-for cause of the death "when and as" it happened.

This form of reasoning is nothing more than proof by tautology. For example, consider Perkins' treatment of the case in which the victim is struck simultaneously by two bullets, each of which would have been instantly fatal by itself: "[He] would not have died when *and as* he did die (by two bullets) had only one been fired. 'One might have caused the result, but in fact both did so.'"¹⁷⁷ The factors believed to be causally relevant (the two bullets) are incorporated into the description of the manner of occurrence of the injury (death by two bullets), and they are then demonstrated to be causally relevant because we cannot construct that precise description without them.

Such an approach could just as easily have included the victim's knife wound, his silk shirt, or the air temperature in the description of the manner of occurrence of the injury. If so included, each of these conditions would also be proven under Perkins' test to be but-for causes of the harm. We do not include them because we do not believe they were causally relevant. But that is precisely the issue the test is supposed to determine. The test "proves" only what has already been decided. It does not tell us how that decision was made or help us to make it in hard cases.¹⁷⁸

2. Detailing the Injury

Becht and Miller accept Perkins' reasoning,¹⁷⁹ but they focus more on the details of the harm itself—the final result—than on the manner of its occurrence.¹⁸⁰ For example, in the case of the two fires, one set by the defendant and each sufficient to destroy the plaintiff's house, Becht and

176. R. PERKINS, CRIMINAL LAW 689 (2d ed. 1969) (emphasis in original).

177. *Id.* (emphasis in original) (citation omitted).

178. Hart and Honoré, although recognizing that prior causal determinations are hidden in this method, seem to approve its use. H.L.A. HART & T. HONORÉ, *supra* note 8, at xli-xlii, 124-25, 235 & n.56, 252-53; *accord* J. FLEMING, *supra* note 4, at 173-74; R. KEETON, *supra* note 104, at 7; Rizzo, *The Imputation Theory of Proximate Cause: An Economic Framework*, 15 GA. L. REV. 1007, 1020 (1981).

179. A. BECHT & F. MILLER, *supra* note 8, at 17; *see also id.* at 210.

180. One of their examples, discussed in the text, refers to being hit by "different parts of the car" (manner of occurrence) as well as to "slightly different injuries" (result), but almost all the other examples refer to different details of the result. However, their suggested approach of using the but-for test at short intervals throughout the causal process, rather than only once at the time the

Miller assert that, using the but-for test applied "in minute detail, it would probably appear that the defendant's fire was a cause, for the positions of the smoke, ashes, and some parts of the ruins might well have been caused by [i.e., not have existed but for] the defendant's fire."¹⁸¹

Again, this assertion begs the question. Under this approach, the defendant's fire was a but-for cause only of the precise "positions of the smoke, ashes, and some parts of the ruins," not of the destruction of the house, for which these precise details are irrelevant. The detailed description will be useful only if the plaintiff wishes to recover for damages that would not have occurred but for the precise position of the smoke and debris—a most unlikely lawsuit.

If being a but-for cause of even the most trivial detail of a harm were sufficient to make the defendant a cause of all the harm, the detailed but-for approach would prove too much. Becht and Miller discuss an example in which an inattentive driver hits a pedestrian who suddenly ran into the path of the driver's car. If the driver had been attentive, he could have swerved a little, but not enough to avoid impact with serious injury. Nevertheless, a detailed application of the but-for test shows that the driver's inattentiveness "was a cause of the injuries actually suffered," since "the injuries would have been at least slightly different and would have been inflicted by different parts of the car if the driver had swerved."¹⁸² Becht and Miller admit that both laypersons and lawyers, using the common knowledge of causation upon which Becht and Miller rely so strongly,¹⁸³ would conclude that the driver's inattentiveness was not a cause in such a case. They suggest that laypersons and lawyers make this mistake by failing to pay careful attention to the precise details.¹⁸⁴ But laypersons and lawyers, even after considering the precise details, would still assert correctly that the inattentiveness was not a cause of the serious injury, although it may have been a (but-for) cause of some details associated with the injury. These details have causal significance only if they contributed to the seriousness of the injury.

Becht and Miller realize that they, rather than the laypersons and lawyers, must give way on this issue. They compromise their usual causal principles by "equating the injuries." They reluctantly disregard

process is completed, is similar to Perkins's detailing of the manner of occurrence. *See id.* at 15-17, 19.

181. *Id.* at 18.

182. *Id.* at 28. Becht and Miller use the phrases "hypothetical inquiry" and "construction of parallel series," but they acknowledge that their approach is a narrow form of the but-for test. The test is applied step-by-step and in minute detail, and it excludes consideration of causes that the actor could not have prevented. *Id.* at 17-21, 32; *see infra* text accompanying notes 211 & 219.

183. A. BECHT & F. MILLER, *supra* note 8, at 9-11; *see infra* text accompanying notes 209-11, 217.

184. A. BECHT & F. MILLER, *supra* note 8, at 28-29.

minor differences and conclude that the inattentiveness was not a cause.¹⁸⁵ Although they do not acknowledge it, they thereby destroy the alleged usefulness of the minute-detail approach. After utilizing the minute-detail approach to support a finding of causation, they then are forced to ignore the details and reject causation by equating the injuries. The concept of "equating the injuries" introduces an inconsistency into their theory that undermines their use of the minute-detail approach to support a finding of causation in the merged-fires case.

3. *Excluding Hypothetical Facts*

Glanville Williams also refers to Perkins' approach of detailing the manner in which the harm occurred, but he proposes a different approach for the preemptive-causation cases. In such cases, facts that did not actually occur must not be "invented" and considered, no matter how likely they were to occur. Thus, the hypothesized but almost certain fact that *P* would have been shot by *C* if he hadn't already been shot by *D* cannot be considered.¹⁸⁶ However, on the very next page, Williams is obliged to depart from this rule to reach the correct causal conclusion in a case of negligent omission.¹⁸⁷ Moreover, such a rule is of no help at all in cases where the preempted condition actually has occurred—for example, when *P* actually was shot by *C* after he already had been shot and killed by *D*. In this type of preemptive-causation case, and in the duplicative-causation cases, Williams abandons the but-for test for actual causation and treats the causal issue as a policy issue.¹⁸⁸

4. *Aggregating Multiple Potential Causes*

The editors of the fifth edition of Prosser's hornbook on tort law attempt to resolve the causal issue in the overdetermined-causation cases by applying the but-for test to the aggregate of potential causes:

When the conduct of two or more actors is so related to an event that their combined conduct, viewed as a whole, is a but-for cause of the event, and application of the but-for rule to them individually would absolve all of them, the conduct of each is a cause in fact of the event.¹⁸⁹ This version of the but-for test, like Perkins's version, works only if the person applying the test already has correctly determined which factors

185. *Id.* at 29-31.

186. Williams, *supra* note 21, at 72.

187. *Id.* at 73.

188. *Id.* at 75-77.

189. PROSSER & KEETON, *supra* note 7, § 41, at 268; see also 2 F. HARPER & F. JAMES, *supra* note 2, § 20.2, at 1110; Carpenter, *supra* note 127, at 406-07; Carpenter, *Concurrent Causation*, 83 U. PA. L. REV. 941, 944-45, 952 (1935); cf. A. BECHT & F. MILLER, *supra* note 8, at 211 (mentioning the aggregate but-for test); Cole, *supra* note 7, at 810-11 (defendant liable if duplicative or preemptive condition has same relation to plaintiff under the compensation rule as defendant does).

are sufficiently "related" to the event to be treated as its causes. On its own, the test cannot distinguish duplicative causes from preempted conditions. Moreover, it would treat totally unrelated conditions as causes.

For example, consider the case in which *C* poisons *P*'s tea and *D* shoots *P* before *P* drinks the tea. Assume also that *B* was climbing Mt. Everest. Clearly, *D*'s shooting *P* was a preemptive cause of *P*'s death, and neither *C*'s poisoning the tea nor *B*'s climbing Mt. Everest was a cause. The proposed test, however, would treat all three activities as causes. In the aggregate they were a but-for cause of *P*'s death, while individually none of them was a but-for cause.

C. The Substantial-Factor Formula

Those theorists and judges who are unwilling to engage in contortions such as the ones illustrated above to rescue the but-for test from its inadequacies usually have turned to the substantial-factor formula, either as a supplement to the but-for test (to apply to the overdetermined-causation cases) or as a complete substitute for it.

The substantial-factor formula was created originally not as a test of actual causation but as a guide for resolving proximate-cause issues. The originator of the formula, Jeremiah Smith, was content with the but-for test as a test of cause-in-fact, with the usual exception for the overdetermined-causation cases.¹⁹⁰ He wanted to devise a practical alternative to the probability or foreseeability tests for determining the proximate-cause limitations on liability, because he believed that those tests were unsound and inconsistent with the results in many cases.¹⁹¹ He proposed the following formula: "Defendant's tort must have been a substantial factor in producing the damage complained of."¹⁹² The accompanying explanation and alternative formulations clearly stated that the defendant's tort could not be a substantial factor unless it satisfied the but-for test (with an exception for simultaneous independent sufficient causes); in addition, it would have to be an appreciable and continuously effective or efficient factor in producing the harm, up to the time of occurrence of the harm.¹⁹³

Thus, the substantial-factor formula was meant to be used as the test of legal (proximate) cause, but also incorporated the but-for test (and its exception) for cause-in-fact. Smith's approach was adopted essentially

190. Smith, *supra* note 106, at 108-09 & n.20, 227, 312, 314 n.36, 316-17 n.41. Smith's exception for overdetermined causation, however, was very narrow. It encompassed only a very small subset of such cases—those in which there are two independent simultaneous tortious causes, each by itself sufficient to produce the harm. *Id.* at 109 n.20, 312, 316-17 n.41.

191. *Id.* at 105, 115-28, 223-52, 308-09.

192. *Id.* at 309.

193. *Id.* at 310-12, 314 n.36.

intact in the original *Restatement of Torts*.¹⁹⁴ It has persisted in the *Restatement (Second) of Torts'* treatment of legal causation and cause-in-fact, despite an attempt by Prosser and others to confine the substantial-factor formula to the question of causation-in-fact.¹⁹⁵

Leon Green first suggested that the substantial-factor formula be applied to the actual-causation issue. He completely rejected the but-for test, since it "take[s] the eye off the ball" by asking what would have happened rather than focusing on what did happen and thereby denies causation in many cases where causal contribution to the injury is obvious.¹⁹⁶ In his view, the substantial-factor formula was the best available alternative. He argued that it should be used without elaboration to pose the issue of causal contribution to the jury in every case, not just in the overdetermined-causation cases.¹⁹⁷ Green's view was accepted by Prosser, who passed it on to several generations of law students and lawyers through the various editions of his hornbook,¹⁹⁸ even after Green himself became disenchanted with the formula.¹⁹⁹

The problem with the substantial-factor formula as a test of actual causation (apart from its complete lack of guidance on what constitutes a "factor") is that the alleged cause must be a *substantial* factor. Thus, paradoxically, at the same time that Green was arguing strenuously and cogently that the issue of causation should be completely divorced from the policy issues camouflaged in proximate-cause analysis, he adopted a formula for actual causation that required the judge or jury to determine not only whether the actor's tortious conduct had contributed to the injury (been a factor), but whether it had contributed *enough* to make the actor responsible—whether, "in the light of all the other factors, the defendant's conduct played an *appreciable* part in the result."²⁰⁰ Green seemed to believe that the resulting inquiry retained its character as a factual, causal inquiry because it allegedly relied on quantitative rather than qualitative measures of relative contribution.²⁰¹

Even if quantitative measures of relative contribution are used, however, the question of limiting liability due to the *extent* of contribution,

194. RESTATEMENT OF TORTS §§ 431-435 (1934).

195. RESTATEMENT (SECOND) OF TORTS §§ 431-433 (1965). Compare W. PROSSER, *supra* note 1, § 42, at 248 (asserts formula confined to cause-in-fact issue) with A. BECHT & F. MILLER, *supra* note 8, at 16, 130-34 (criticizes that assertion).

196. Green, *supra* note 8, at 556; see also L. GREEN, JUDGE AND JURY 192 (1930); L. GREEN, *supra* note 2, at 167-70; Green, *supra* note 158, at 604-05; Green, *supra* note 8, at 557-59.

197. L. GREEN, JUDGE AND JURY 190-95 (1930); L. GREEN, *supra* note 2, at 137, 140, 181-85; Green, *supra* note 158, at 603-07.

198. E.g., W. PROSSER, *supra* note 1, § 41, at 239-40; *id.* § 46, at 321, 323-24 (1st ed. 1941).

199. See *infra* text accompanying notes 206-08.

200. L. GREEN, *supra* note 2, at 134 (emphasis in original); see *id.* at 134-37, 141.

201. *Id.* at 140-41; see L. GREEN, JUDGE AND JURY 192-93 (1930); L. GREEN, *supra* note 2, at 84, 122.

rather than due to the absence of *any* contribution, is clearly a proximate-cause issue of policy or principle, rather than an issue of actual causation (contribution to the injury). Moreover, the substantial-factor formula inherently invites consideration of qualitative measures as well as quantitative ones.²⁰²

Green himself, immediately after emphasizing the quantitative nature of the inquiry, used qualitative considerations when discussing a series of cases involving multiple causes. For example, he asserted that it is "perhaps clear" that the hole poked by *D*'s tug in *P*'s barge, three feet above the water line, was not an appreciable factor in the sinking of the barge when compared with the continued loading of the barge which brought the hole below the waterline, since those who were loading the barge were fully conscious of what had occurred but made no effort to repair the hole. Similarly, he suggested that there would be little possibility of a finding that *D*'s driving over a small flat box, which unbeknown to him contained a highly explosive substance, was a substantial factor in producing the resulting explosion. And, when discussing the cases in which *P* is negligently or intentionally pushed by *X* into a hole negligently maintained by *D*, Green stated that "what would be considered appreciable in one case might not in another. So, as compared with *X*'s negligence, [D's] conduct might be considered appreciable, while as compared with *X*'s intended violence, [D's] conduct might not be so considered."²⁰³

In each of these cases, the defendant's tortious act was a necessary condition for the occurrence of the injury. When an act (omission, condition) was a *necessary* condition for the occurrence of an injury, it is unnatural to deny that the act was, as a matter of *fact*, not only a factor in producing the injury, but also a substantial or appreciable factor. To state otherwise is to make a noncausal, nonfactual policy judgment about responsibility for the injury.

Prosser and, eventually, Green seem to have recognized the emptiness of the undefined substantial-factor formula and the danger of its being used to introduce proximate-cause issues into the actual-causation inquiry. Prosser, in the various editions of his hornbook, concluded his brief discussion of the formula by equating it with the but-for test, with an exception for the overdetermined-causation cases,²⁰⁴ and in his articles he relied primarily on the but-for test.²⁰⁵

202. A. BECHT & F. MILLER, *supra* note 8, at 131-34; H.L.A. HART & T. HONORÉ, *supra* note 8, at 97, 293-94; C. MORRIS & C.R. MORRIS, *supra* note 4, at 174-75; Malone, *supra* note 6, at 88-97.

203. L. GREEN, *supra* note 2, at 153-58.

204. E.g., W. PROSSER, *supra* note 1, § 41, at 240 & n.33. The fifth edition of Prosser's hornbook, edited by others, recognizes the policy judgments inherent in the substantial-factor formula. PROSSER & KEETON, *supra* note 7, § 41, at 266-69.

205. E.g., Prosser, *Proximate Cause in California*, 38 CALIF. L. REV. 369, 375-78 (1950).

In his later writings, Green abandoned his earlier substantive interpretation of the substantial-factor formula. He treated it as a meaningless label for the judge to attach to her determination on the actual-causation issue or to use as a hortatory device "to caution a jury to weigh the evidence carefully."²⁰⁶ He also criticized the view that the causal inquiry involves not only the question of contribution to the injury, but also the question of how significant the contribution was, without noting that he previously had been the major proponent of this view.²⁰⁷ However, he continued to reject the but-for test and to insist that the issue of causal contribution be submitted to the jury without any elaboration or guidance.²⁰⁸

D. Undefined, Directly Observable Causal Contribution

Lacking any satisfactory test, Green and a few other writers have treated causation as an undefinable and irreducible factual relation between events that can be directly perceived or inferred without explicitly or implicitly resorting to any generalization, definition, or test.²⁰⁹ Most of these writers do not attempt to explain how this concept would be applied in the various types of causation cases. However, there is one notable exception: Becht and Miller's book, *The Test of Factual Causation*.²¹⁰

Becht and Miller's book is a comprehensive attempt to demonstrate how the concept of undefined, directly observable causal contribution is applied in tort analysis. They argue that we determine whether a causal relation exists between an act and an injury by breaking down the sequence of events to the appropriate level of detail to see if we can "perceive" a causal connection. If we can, the act is deemed to be a cause of the injury even if the injury can also be traced back to another act or event through a different causal sequence as in the merged-fires case.²¹¹ However, when Becht and Miller attempt to show how this approach would work in the myriad variety of actual-causation cases, they are forced to rely on the but-for test in an increasingly wide range of cases.

Initially, they are forced to use the but-for test to handle causal inquiries involving omissions. By definition an omission is a nonevent—something which did not happen—which only rarely will trigger an

206. Green, *supra* note 8, at 554.

207. *Id.* at 555; *see id.* at 557.

208. *Id.* at 548, 553-58.

209. A. BECHT & F. MILLER, *supra* note 8, at 9-10, 163-64; L. GREEN, *supra* note 2, at 132-33, 135, 137, 139-40, 183-85; Green, *supra* note 8, at 549, 553 & n.22, 554 n.25, 556, 560; *see* Borgo, *supra* note 11, at 434-38; Epstein, *supra* note 13, at 161-63, 165-67; *cf.* Williams, *supra* note 21, at 66-68 (causal generalizations not involved in causal inquiry in the legal context).

210. A. BECHT & F. MILLER, *supra* note 8.

211. *Id.* at 9-11, 14-15, 18.

actual causal sequence that can be directly perceived or traced. Instead, an omission is almost always a "negative" cause of an injury: it represents a missed opportunity to *prevent* the injury. That is, an omission is a cause of an injury only if the omitted act would have been part of a hypothetical causal sequence that would have prevented, terminated, or deflected the different causal sequence that actually occurred and produced the injury. Thus, in order to determine whether an omission was a cause of an injury, it is necessary to conduct a hypothetical inquiry. The omitted act must be hypothetically supplied, and a hypothetical causal sequence (which Becht and Miller call a "parallel series") must be constructed and traced to determine whether it would have prevented the occurrence of the injury.²¹²

The hypothetical inquiry (but-for test) also must be used whenever an actual ("positive") causal sequence becomes a negative cause by blocking or terminating some preventive effort—for example, by blocking a fire engine or cutting off its supply of water.²¹³ This inquiry also must be employed to determine whether the excess (negligent) portion of a certain speed or weight, or the difference between being in a safe position or an unsafe (negligent) position, contributed to the injury.²¹⁴ Indeed, as Becht and Miller reluctantly admit, the hypothetical inquiry must be used whenever the negligent aspect of an act does not encompass the act as a whole. For example, in the excess-speed case, the act is driving at a certain speed. We observe the car at that speed hit the child. The negligent aspect of the speeding is the excess speed, which Becht and Miller note "is not an event but a concept. Like an omission, it did not happen, and it cannot be a simple [perceived] cause of events in the world."²¹⁵

Becht and Miller recognize the possible criticism that they are implicitly using the hypothetical inquiry with all negligent acts, rather than just with negligent acts that have a nonnegligent aspect, "but are unconscious of [doing so] when the [hypothetical inquiry] shows clearly that the negligent act was a cause of the harm."²¹⁶ They admit that they cannot disprove this contention, but they continue to claim that they are using direct perception of causation in such cases. However, they have

212. *Id.* at 21-24. Becht and Miller assert that an omission can only be a negative cause—that it can only contribute by failing to prevent or block the positive causal sequence that actually produces the result, rather than by itself triggering the actual causal sequence that produces the result. *Id.* at 170-71. But an omission can trigger an actual causal sequence. For example, an omission to salute causes a soldier to be disciplined, or an omission to put money in the parking meter causes the driver's car to be ticketed. The mental perception of the omission provides the connecting link.

213. *Id.* at 41-42.

214. *Id.* at 52-63.

215. *Id.* at 54; see also *id.* at 33, 52, 63, 65, 140.

216. *Id.* at 66.

some difficulty with a case in which the defendant negligently welds the roof of a car, leaving a sharp edge, and the driver's head is crushed against the roof at the weld point when the roof is smashed in during an accident. They steadfastly assert that the case is one in which direct perception or inference, unaided by any hypothetical inquiry, shows a lack of causation by the sharp weld, just as direct perception or inference shows that the color of the roof was not a cause of the injury.²¹⁷

The direct-perception argument relies on the notion that there are observable and describable causal qualities or forces in objects or events. Becht and Miller eventually concede that this notion is naive as compared with the Humean account, which is based on causal generalizations.²¹⁸ Even if the direct-perception notion were creditable, it is difficult to imagine how the lack of causation by the sharp weld could be perceived in the circumstances given. It is also difficult to understand what is meant by "direct inference." An inference is based on some concept of how things *generally* happen—that is, on causal generalizations. Causal generalizations incorporate the belief that the cause is in some sense necessary for the occurrence of the consequence. The but-for test is simply the means by which we determine whether this element of necessity exists in the particular case.

Becht and Miller argue that, even if it is true that they are explicitly or implicitly using the but-for test in every case, they are using a less restrictive version. Unlike the usual version, they compare what happened with what would have happened in as minute detail as possible. Moreover, in those cases in which there was more than one positive causal sequence that led to the injury, they only require that the defendant's act or omission have been necessary for the continuation of the positive causal sequence with which it was associated, rather than requiring that it have been necessary for the injury itself.²¹⁹

The first distinction—their use of minute detail—is nullified by their tactic of "equating the injuries" to avoid finding but-for causation whenever there is a slight difference between the actual process or result and the hypothesized process and result.²²⁰

The second distinction, however, is a real one. It enables them to find that causation exists in many of the overdetermined-causation cases. For example, in the merged-fires case the defendant's tortious conduct will be treated as a cause of the destruction of the house if it was a but-for cause of one of the fires, even though the house would have been burnt

217. *Id.* at 66-67; *see also id.* at 92, 139.

218. *Id.* at 163-64; *see also infra* text accompanying notes 228-31; T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 80-84, 107-17.

219. A. BECHT & F. MILLER, *supra* note 8, at 65-67; *see also id.* at 19, 31, 64-65.

220. *E.g., id.* at 28-31, 55, 58-59, 62, 64, 70-71, 91; *see supra* text accompanying notes 179-85.

down by the other fire anyway. The defendant's tortious conduct was a but-for cause of one of the duplicative causal sequences, and that is sufficient to establish the causal connection to the destruction of the house.²²¹

Becht and Miller's causal-contribution analysis ends up being very similar to the "necessary element of a sufficient set" test.²²² However, it is not exactly equivalent. Indeed, it produces anomalous results precisely to the extent that it varies from the latter test. Becht and Miller's approach requires that the tortious act or omission have been a necessary condition for one of the positive (actual) causal sequences that contributed to the injury. They are therefore forced to deny that causation exists when the overdetermined-causation problem results from duplicative or preempted conditions that affect the same positive causal sequence, rather than from the existence of multiple positive causal sequences.

For example, they discuss the case in which *C* negligently fails to discover and repair defective brakes in a car before renting the car to *D*, who, unaware of the defective condition of the brakes, negligently fails to try to use them to avoid hitting *P*. *P* is hit, but he would not have been hit if the brakes had been working properly and *D* had used them. *C*'s omission (the failure to inspect and repair the brakes) was not a necessary condition for the occurrence of the positive causal sequence that produced the injury, since *D* made no attempt to use the brakes. Similarly, given *C*'s omission, *D*'s omission (the failure to try to use the brakes) was not a necessary condition either. Becht and Miller unhappily but firmly conclude that neither omission was a cause of the injury, but would hold both *C* and *D* liable to avoid a "morally indefensible" and "inexcusable" result.²²³

The same problem occurs with acts. Assume that a cable has a maximum safe load capacity of one ton when it is in good condition. *C* negligently weakens the cable (for example, by cutting a few of its strands) so

221. A. BECHT & F. MILLER, *supra* note 8, at 32.

222. Becht and Miller refer to and approve Hart and Honoré's "sufficient set" approach to causation, but they avoid mentioning or discussing the "necessary element" aspect and in fact erroneously imply that Hart and Honoré do not believe that the notion of necessity is relevant to causation. *Id.* at 202-04, 222. For a discussion of Hart and Honoré's position, see *infra* text accompanying notes 228-32, 235-38, 252-60, 282-83.

223. A. BECHT & F. MILLER, *supra* note 8, at 95-98; see also *id.* at 82 n.116, 210-11. Prosser reaches a similar conclusion:

It may be said with some confidence that if any such case is considered, both parties will be held liable; but the theory of liability is not so clear. Perhaps the best guess is that each, by his negligence, has deprived the plaintiff of a cause of action against the other, and so should be liable.

W. PROSSER, *supra* note 1, § 41, at 239-40 n.25; accord J. FLEMING, *supra* note 4, at 174 n.19. Green's treatment of the case is discussed *supra* text accompanying notes 108-15. For the resolution of the case under the "necessary element of a sufficient set" test, see *infra* text accompanying notes 274-76.

that it will now break if a one-ton load is applied. *D* negligently overloads the cable with a two-ton load, which would cause it to break even if it were in good condition. The cable breaks at the weakened point, and the load falls on and injures *P*. Since the negligent acts of *C* and *D* are both part of the same positive causal sequence that produced the injury, but neither was a necessary condition for the occurrence of the sequence, neither is a cause of the injury under Becht and Miller's approach.²²⁴

Becht and Miller's book contains illuminating discussions on many points—for example, the distinction between acts and omissions and the need to focus on the tortious aspect of the actor's conduct. Perhaps its most valuable contribution, however, is an unintended one: the demonstration of the unavoidability of the but-for (necessary condition) test in causal analysis. This unavoidability is most clear when the causal inquiry focuses on omissions or the tortious aspect of acts, but it is implicit in all causal inquiries.²²⁵

E. The NESS (Necessary Element of a Sufficient Set) Test

In this section, I turn at last to a test for causal contribution that is applicable to the entire spectrum of causation cases. This test incorporates the indispensable notion of necessity, but subordinates it to the notion of sufficiency. I call this test the NESS (Necessary Element of a Sufficient Set) test.

The NESS test was first suggested by H.L.A. Hart and Tony Honoré. However, their brief exposition of this test was overshadowed and distorted by their primary emphasis on proximate-cause issues.²²⁶ The test has received very little notice and even less acceptance in the subsequent legal literature.²²⁷ One of the principal purposes of this Article

224. A. BECHT & F. MILLER, *supra* note 8, at 57-58 & n.85. See *infra* text following notes 245 & 250 for the disposition of the case under the "necessary element of a sufficient set" test.

225. This point, at least with respect to omissions and the tortious aspect of acts, was not missed by Green, who was a lifelong critic of the but-for test. In his earlier writings, he had focused the causal inquiry on the negligent aspect of the actor's conduct. After reading Becht and Miller's book, however, he insisted that the "orthodox analysis" limits the causal relation issue to the question of whether the actor's conduct as a whole contributed to the injury. Green, *supra* note 8, at 546, 549-51. He also claimed that there is no need to distinguish omissions from acts since "omissions to act are merely incidents in a longer line of affirmative conduct and are easily resolved into it." *Id.* at 546; see *id.* at 546-47. We have already seen that this approach produces anomalous results unless the tortious-aspect causation requirement is smuggled back in as part of the duty or proximate-cause analysis. This is, in fact, what Green and his followers have done, with mixed success. See *supra* text accompanying notes 111-26.

226. H.L.A. HART & A. HONORÉ, *supra* note 10, at 105-10, 116-19, 122, 216-29; H.L.A. HART & T. HONORÉ, *supra* note 8, at 111-17, 122-25, 128-29, 235-53.

227. Fleming incorporated much of Hart and Honoré's exposition, including its deficient aspects, into the second edition of his treatise. J. FLEMING, THE LAW OF TORTS 178-81 & n.1 (2d ed. 1961). The exposition has been retained with slight modifications in subsequent editions. E.g., J. FLEMING, *supra* note 4, at 170-71, 173-74. Professors Cole, Becht, and Miller omit the necessary-condition element and treat the test as a "sufficient set" approach. A. BECHT & F. MILLER, *supra*

cle is to draw renewed attention to the NESS test, by demonstrating its fundamental identity with the basic concept of causation and its utility in resolving the cases that have proved troublesome for several generations of tort scholars.

1. *The Philosophic Basis of the NESS Test*

The NESS test captures the essential meaning of the concept of causation. This meaning was first articulated by the philosopher David Hume. Hume rejected the earlier notion that we acquire causal knowledge through direct sensory perception of causal qualities or forces inherent in objects or events. Instead, he insisted that we only observe certain successions of events, more or less frequently repeated. From these observations, we inductively derive the belief that certain antecedent events are not only always conjoined with, but also are sufficient for the occurrence of, certain subsequent events. That is, we form a belief in more or less well-grounded causal laws or generalizations. Causal knowledge or belief also can be based on reasoning by analogy from such observations or can be acquired through education, which passes on the causal knowledge of others. Any singular causal statement about a particular occurrence is not a reporting of direct sensory perception of causal forces, but rather an assertion of the belief that the occurrence instantiates one or more causal laws or generalizations.²²⁸

A fully specified causal law or generalization would state an invariable connection between the cause and the consequence: given the actual existence of the fully specified set of antecedent conditions, the consequence must follow. In other words, the fully specified set of antecedent conditions is *sufficient* for the occurrence of the consequence. In the typical singular causal statement, the causal assertion includes, explicitly or implicitly, only a few of the antecedent conditions but nevertheless asserts that they were part of an incompletely specified (and incompletely understood or known) set of actual conditions that was sufficient for the occurrence of the consequence.²²⁹

note 8, at 202-04, 222, *discussed supra* note 222; Cole, *supra* note 7, at 485, 486 n.80, 488, 495, 505 & n.133, 769, 770 n.16. Professors Fraser and Howarth also mention the NESS test, but attribute it to Professor Papineau rather than to Hart and Honoré and reject it, after misapplying it, in favor of a probabilistic increased-risk "causal" test. Fraser & Howarth, *More concern for cause*, 4 *LEGAL STUD.* 131, 133 & n.19, 135-38, 140-42, 145-56 (1984). Similarly, Mario Rizzo converts the INUS (Insufficient but Necessary element of an Unnecessary but Sufficient set) test, which is similar to the NESS test and was elaborated by Mackie subsequent to Hart and Honoré's book, into a probabilistic-linkage test. Rizzo, *supra* note 178, at 1009-16; *see infra* text accompanying note 234.

228. T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 4-11, 80-81, 87-88, 91-92, 139-43; H.L.A. HART & T. HONORÉ, *supra* note 8, at 10-11, 14-15.

229. T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 23-24, 84-87, 92-97, 114-15, 132, 136-37, 141; H.L.A. HART & T. HONORÉ, *supra* note 8, at 10-11, 44-49; Mackie, *supra* note 21, at 23-24, 27-32, 35-38.

Hume was primarily interested in elaborating and defending this "regularity" account of causation and causal knowledge, which insists that the meaning of causation is to be found in causal laws or generalizations rather than in some notion of causal qualities or forces. However, while pursuing this primary objective, he occasionally noted that one must be careful when constructing a causal law or generalization to distinguish the causally relevant antecedent conditions from the causally irrelevant antecedent conditions. This differentiation is necessary to insure that the set of jointly sufficient antecedent conditions includes only those that are indeed invariably connected with the consequence. Thus, the antecedent conditions must be restricted to those that are *necessary* for the sufficiency of the set.²³⁰

Hume also maintained that a certain consequence is always produced by the same cause—that is, that there is a unique sufficient set of antecedent conditions that always must be present to produce a particular consequence. Hume's successor, John Stuart Mill, disagreed. He asserted that there may be a plurality of potential causes for any consequence. For example, death may be caused by poison on one occasion, by a bullet wound on another, by fire on yet another, and so forth. Hume's defenders have argued that such seemingly different causes may upon further investigation turn out to be overly broad specifications of a single common cause, or that the seemingly identical consequence may turn out to be distinct consequences. However, they also admit that ordinary experience provides strong support for the plurality-of-potential-causes theory. Indeed, the plurality theory has become part of the dominant "regularity" account of the meaning of causation.²³¹

As Hart and Honoré noted in 1959, the "necessary element of a sufficient set" (NESS) test of causal contribution follows directly from this dominant "regularity" account of the meaning of causation.²³² The essence of the concept of causation under this philosophic account is that *a particular condition was a cause of (condition contributing to) a specific consequence if and only if it was a necessary element of a set of antecedent actual conditions that was sufficient for the occurrence of the consequence.* (Note that the phrase "a set" permits a plurality of sufficient sets.) This is the more precise, extended statement of the NESS test.

Surprisingly, however, this notion of a contributing condition appears only infrequently in the philosophic literature on causation. The literature continues to be characterized by arguments based on necessary

230. T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 94-96, 114-17.

231. *Id.* at 23, 88-91; H.L.A. HART & T. HONORÉ, *supra* note 8, at 19-20; *see id.* at xxxix-xlii.

232. H.L.A. HART & T. HONORÉ, *supra* note 8, at 111-18. Sosa incorrectly characterizes Hart and Honoré's account as a simple sufficient-condition account. Sosa, *supra* note 170, at 1; *see A. BECHT & F. MILLER, supra* note 8, at 202-04, 222, *discussed supra* note 222.

conditions, sufficient conditions, or necessary and sufficient conditions.²³³ After Hart and Honoré's book was published, several nonlegal philosophers articulated NESS-like notions of causal contribution. In each instance, however, they burdened the NESS test with the additional requirement that the sufficient set be necessary for the occurrence of the consequence in the particular circumstances. This requirement converts the NESS test to a necessary and sufficient condition test, which cannot accommodate the duplicative-causation cases (for example, the merged-fires cases).²³⁴

Hart and Honoré do not make this mistake, because they recognize that there may be a plurality of causes in a particular instance. However, their version of the NESS test is also deficient in several respects. First, they apparently require that a duplicative or preemptive cause be independently "sufficient" for the occurrence of the injury. That is, it must be sufficient in conjunction with the "background" conditions, but excluding the duplicated or preempted conditions. Second, they submerge the critical distinction between duplicative and preemptive causation by constructing an overlapping typology of overdetermined-causation cases. Finally, they confuse the factual issue of causal contribution with the issues of policy or principle involved in determining ultimate liability in certain types of cases.

2. *The Duplicative-Causation Cases*

As elaborated, Hart and Honoré's NESS test is only slightly less restrictive than the *Restatement*'s substantial-factor exception to the but-for test, which only applies when each of two actively operating factors was sufficient by itself for the occurrence of the injury.²³⁵ Hart and Honoré do not require that each of the factors have been actively operating, but they seem to require that each have been sufficient by itself for the occurrence of the injury.²³⁶ For example, in the merged-fires cases, they assume that each fire would have been sufficient by itself for the destruction of the plaintiff's house. The NESS test then confirms causal contribution by each fire. Each fire was necessary for the sufficiency of a set of actual antecedent conditions that did not include the other fire.²³⁷

The requirement that each factor have been sufficient by itself (when combined with the background conditions) is too restrictive and is not a part of the basic concept of causation that is reflected in the NESS test.

233. Sosa, *supra* note 170, at 1-3.

234. Mackie, *supra* note 21, at 16-21, 25-27 (also discussing Marc-Wogau's and Scriven's similar accounts); see T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 212-13; Sosa, *supra* note 170, at 3-4.

235. RESTatement (SECOND) OF TORTS § 432(2) (1965).

236. H.L.A. HART & T. HONORÉ, *supra* note 8, at 123-24, 125, 206-07, 235-39, 245, 249.

237. *Id.* at 122-23, 206, 235, 237-39.

Moreover, the independent-sufficiency requirement is not followed by the courts. For example, in two duplicative-causation cases involving merged fires and noisy motorcycles, respectively, the courts did not require the plaintiff to prove the independent sufficiency of each contributing factor, but rather required him to prove only that each factor contributed to the injury.²³⁸ Similarly, in the pollution cases, the courts have allowed the plaintiff to recover from each defendant who contributed to the pollution that caused the injury, even though none of the defendants' individual contributions was either necessary or sufficient by itself for the occurrence of the injury.²³⁹

Malone argues that the courts' departure from the but-for test in the overdetermined-causation cases is further evidence of the policy-dependent nature of the causal inquiry. He contends that the courts are willing to replace the but-for test with the substantial-factor formula in cases like the merged-fires cases because of the important policies that underlie "the well-established rules that prohibit the setting into motion of a destructive force."²⁴⁰

But, contrary to Malone's assertion, the policies or principles at work in these cases are not those underlying the particular rule that was violated. They are the ones underlying the actual causation requirement itself. As Malone himself notes, the courts feel impelled to depart from the but-for test in these overdetermined-causation cases because, although the but-for test is not satisfied, it is clear that the defendant's tortious conduct²⁴¹ contributed to the injury: "Our senses have told us that he *did* participate. . . . In the language of the layman, the defendant's fire 'had something to do with' the burning of the plaintiff's property."²⁴² Just as Newtonian mechanics serves as an adequate substitute for the more accurate and comprehensive theories of relativity and quantum mechanics in ordinary physical situations, the but-for test serves as an adequate substitute for the NESS test in ordinary causal situations. In each context, however, the substitute must give way to the more accurate and comprehensive concept when the situation is more subtle and complex.

In the pollution cases, the NESS test confirms that each defendant's

238. *Corey v. Havener*, 182 Mass. 250, 65 N.E. 69 (1902) (noisy motorcycles); *Anderson v. Minneapolis, St. P. & S. Ste. M. Ry.*, 146 Minn. 430, 179 N.W. 45 (1920) (merged fires). The cases are discussed in Malone, *supra* note 6, at 90-91.

239. E.g., *Michie v. Great Lakes Steel Div.*, 495 F.2d 213 (6th Cir.), cert. denied, 419 U.S. 997 (1974); *Warren v. Parkhurst*, 45 Misc. 466, 92 N.Y.S. 725 (1904), aff'd, 105 A.D. 239, 93 N.Y.S. 1009 (1905), aff'd, 186 N.Y. 45, 78 N.E. 579 (1906); see J. FLEMING, *supra* note 4, at 176.

240. Malone, *supra* note 6, at 89.

241. The term "tortious conduct" should always be interpreted to mean the tortious aspect of the actor's conduct. See *supra* text accompanying notes 135-45.

242. Malone, *supra* note 6, at 89 (emphasis in original).

pollution contributed to the injury, even though it was neither necessary nor independently sufficient for the injury. For example, assume that five units of pollution were necessary and sufficient for the injury and that each of seven defendants discharged one unit of pollution. Each defendant can truthfully say that its one unit was neither necessary nor independently sufficient for the injury. But each defendant's one unit was necessary for the sufficiency of a set of actual antecedent conditions that included only four of the other units, and the sufficiency of this particular set of actual antecedent conditions was not affected by the existence of two additional duplicative units.²⁴³

A similar causal situation exists even if one defendant discharges five units of pollution and a second defendant discharges two units. The two units still mix with the five units to produce the injurious seven units. More rigorously, the two units were necessary for the sufficiency of a set of actual antecedent conditions that included only three of the first defendant's five units, a set whose sufficiency was not affected by the existence of two additional duplicative units also provided by the first defendant.

The analysis of the merged-fires cases is analogous. Thus, if any two of three fires were sufficient for the injury, but none by itself was sufficient, each was a cause of the injury since each was necessary for the sufficiency of a set of actual antecedent conditions that included only one of the other fires. The same causal situation exists even if there were only two fires, one of which was independently sufficient and the other of which was not. The first fire was clearly a cause, since it was independently sufficient. But the second fire also was a cause. It was necessary for the sufficiency of a set of actual antecedent conditions which included another fire (the first) that was "*at least* large enough to be sufficient for the injury if it merged with a fire the size of the second fire." The sufficiency of this set is not affected by the fact that the first fire was so large that it would have been sufficient by itself.

The wording of the quoted condition, "*at least* large enough," is not a verbal gimmick. The condition is an actual one that existed on the particular occasion. It describes a certain factual situation, as much as the condition in the pollution case that referred to only three of the first defendant's five units of pollution. In the former case, the size of the first

243. I assume that the injury was not accelerated or aggravated by the extra units of pollution. If it was, causal contribution would be even clearer. I also assume that the units of pollution arrived simultaneously at the site of the injury. Obviously, if five units arrived before the other two and produced the injury before the other two arrived, the first five units were causes of the injury and the last two were not. Their potential effects were preempted by the effects of the first five. See *infra* text accompanying notes 247-51.

fire is broken down into portions; in the latter case, the amount of the first defendant's pollution is broken down into portions.

The NESS test would attribute causal status to a very small fire that merged with an overwhelming large fire. The very small fire was a duplicative cause of any resulting injury. Whether the person who tortiously caused the very small fire should be held liable for any or all of the resulting injury is an issue of policy or principle that comes under the heading of damages.²⁴⁴

The same analysis applies to the weakened-cable hypothetical.²⁴⁵ In that example, a cable with a maximum safe load capacity of one ton was weakened by *C*, who negligently cut a few of its strands so that it would break if subjected to a one-ton load. *D* then negligently put a two-ton load on the cable, which would have caused it to break even if the cable were in good condition, and the cable broke at the weakened point. *C*'s weakening of the cable was necessary for the sufficiency of a set of actual antecedent conditions which included a load of at least one ton, and the sufficiency of this set was not affected by the fact that the load was more than one ton. Indeed, the causal contribution of *C*'s weakening of the cable is evidenced by the fact that the cable broke at the weakened point.²⁴⁶ Similarly, *D*'s overloading of the cable was necessary for the sufficiency of a set of actual antecedent conditions that included a cable with a load capacity of at most one ton, and the sufficiency of this set was not affected by the fact that the cable had a capacity of less than one ton as a result of the weakening. *C*'s weakening of the cable and *D*'s overloading of it are duplicative causes of the cable's breaking and the resulting injury to the plaintiff.

3. The Preemptive-Causation Cases

So far, the discussion has concentrated on the duplicative-causation cases. I turn now to the preemptive-causation cases. Hart and Honoré mention situations in which the defendant asserts that someone else would have shot the plaintiff or stolen his property if the defendant had not done so.²⁴⁷ The defendant's argument would succeed under a literal application of the *Restatement*, which adheres to the but-for test unless the alternative factor was "actively operating."²⁴⁸ But, under the NESS test, the defendant's tortious act clearly was a cause of the injury, since it

244. See *infra* text accompanying notes 258-73.

245. See *supra* text accompanying note 224.

246. Of course, if the cable broke at another point, *C*'s weakening of the cable would not be a necessary element of any sufficient set of actual antecedent conditions. See *infra* text following note 250.

247. H.L.A. HART & T. HONORÉ, *supra* note 8, at 124-25, 249-50.

248. RESTATEMENT (SECOND) OF TORTS § 432(2) (1965); see *supra* text accompanying note 235.

was a necessary element of the set of actual antecedent conditions that was sufficient for the injury. Under this test, moreover, only the defendant's actions would be deemed a cause. The potential actions of others that did not in fact occur could not be a part of any set of *actual* antecedent conditions that was sufficient for the injury.

Preemptive causation also is present in many cases where the alternative factor actually did occur. For example, in my initial illustration where *D* shot and killed *P* just as *P* was about to drink a cup of tea that had been poisoned by *C*,²⁴⁹ *D*'s shot was necessary for the sufficiency of a set of actual antecedent conditions that did not include the poisoned tea. Conversely, *C*'s poisoning of the tea was not a necessary element of any sufficient set of actual antecedent conditions. A set that included the poisoned tea but not the shooting would be sufficient only if *P* actually drank the tea, but this was not an actual condition. The shooting preempted the potential causal effect of the poisoned tea.

Moreover, even if *P* actually had drunk the poisoned tea, *C*'s poisoning of the tea still would not be a cause of *P*'s death if the poison did not work instantaneously but the shot did. The poisoned tea would be a cause of *P*'s death only if *P* drank the tea and *was alive when the poison took effect*. That is, a set of actual antecedent conditions sufficient to cause *P*'s death must include poisoning of the tea, *P*'s drinking of the poisoned tea, and *P*'s being alive when the poison takes effect. Although the first two conditions actually existed, the third did not. *D*'s shooting *P* prevented it from occurring. Thus, there is no sufficient set of *actual* antecedent conditions that includes *C*'s poisoning of the tea as a necessary element. Consequently, *C*'s poisoning of the tea fails the NESS test. It did not contribute to *P*'s death.²⁵⁰

As the last example illustrates, a necessary condition for the sufficiency of any set of actual antecedent conditions is that the injury not have occurred already as a result of other actual conditions outside the set. The determination of whether this condition existed, as with all the other conditions, is an empirical judgment. For example, in the weakened-cable hypothetical, *C*'s weakening of the cable was considered to be a duplicative cause when the cable broke at the weakened point. If the cable had broken instead at some other point, our empirical evaluation would be that the potential causal effect of *C*'s weakening of the cable was preempted by the existence of this other weaker point in the cable. This other weaker point prevented the cable from staying together long enough for *C*'s weakened point to have any effect. In those circumstances, *C*'s weakening of the cable would not be a necessary element of any sufficient set of actual antecedent conditions.

249. See *supra* text following note 166.

250. See H.L.A. HART & T. HONORÉ, *supra* note 8, at 124.

The key to the overdetermined-causation cases, therefore, is the distinction between duplicative and preemptive causation. In each case, an empirical judgment must be made: was the tortious aspect of the defendant's conduct a necessary element in a set of actual antecedent conditions that was sufficient for the occurrence of the injury, or was it not a necessary member of any actually sufficient set because its potential effects were preempted by other actual conditions outside the set?

In the merged-fires cases or the noisy-motorcycles case, we believe that each factor duplicates or reinforces the effects of the other factor, rather than preempting those effects—that is, that there is more than one sufficient cause. In the shooting case, however, we believe that there was only one sufficient cause (the shooting), which preempted the effects of the other potential cause (the poisoning) and thereby prevented it from being a necessary member of an actually sufficient set. We would believe otherwise if medical experts testified that the effects of the poison combined with the effects of the shooting to cause the victim's death, although either alone would have been sufficient.²⁵¹ The critical issue is whether one factor completely preempted the potential effects of the other factor, so that there was only one sufficient (preemptive) cause in the actual circumstances, or whether each factor was a necessary part of a different (but overlapping) actually sufficient set, in which case there was more than one sufficient (duplicative) cause.

Hart and Honoré submerge this critical issue in their discussion of the overdetermined-causation cases. They construct a typology that first divides the cases into those involving "additional causes" and those involving "alternative causes." The alternative-cause category is meant to include those cases in which there was a potential alternative cause which never occurred. Examples include the cases where the defendant asserts that someone else would have shot or robbed the plaintiff if he had not.²⁵² However, Hart and Honoré also include a bridge-delay case in which there was an *actual* second bridge that would have delayed the plaintiff's boat even if the first bridge had not.²⁵³ It is misleading to speak of these cases as alternative-cause cases. As Hart and Honoré clearly indicate, the alternative cause was not a cause. It was a preempted potential cause, best referred to as a "preempted condition." The category is a subset of the preemptive-causation cases.

Their additional-cause category is subdivided into three subcatego-

251. Thus, the court's holding in *State v. Seates*, 50 N.C. (5 Jones) 420 (1858), that the defendant could not be held liable, even though medical witnesses testified that his burning of a child was the "primary cause" of the child's death, if an independent blow by someone else hastened the death, is clearly erroneous unless the blow was an immediate, preemptive cause of the death. The case is discussed in H.L.A. HART & T. HONORÉ, *supra* note 8, at 242.

252. H.L.A. HART & T. HONORÉ, *supra* note 8, at xl, 124-25, 207, 249-50.

253. *Id.* at 250-51.

ries. The "combinatory" or "reinforcing" cause subcategory, however, is the only one that actually involves additional (duplicative) causes.²⁵⁴ The other two subcategories are subsets of the preemptive-causation cases, where again it is misleading to refer to the preempted condition as an additional cause. The "neutralizing" cause subcategory covers situations in which the effective cause prevents an alternative active causal process from being completed. For example, a fire approaching a house is quenched by a flood which destroys the house, or a poison victim is killed by a bullet before the poison kills him.²⁵⁵ The "overtaking" cause subcategory originally was limited to situations in which a subsequent actual injury would have caused the same damage as the initial injury: for example, two successive injuries, each of which would have been sufficient to disable the victim totally.²⁵⁶ However, the subcategory has been expanded in the second edition to encompass situations in which the effective cause frustrates an active causal process, so that the subcategory now overlaps the "neutralizing" cause subcategory.²⁵⁷

Perhaps as a result of their confusing typology, Hart and Honoré lose sight of the basic concept of causation embodied in the NESS test. In the double-bridge and successive-injury cases, for example, the defendant's tortious conduct was a (preemptive) cause of the boat's delay in the double-bridge case and a (preemptive) cause of the victim's loss of a limb and resulting total disability in the successive-injury case. But Hart and Honoré claim that the tortious conduct was not a cause of the victim's economic loss in either case, since the victim would have suffered the economic loss anyway as a result of the second bridge or the second injury, respectively.²⁵⁸

This causal argument, however, is based on the but-for test rather than the NESS test. In each case, the defendant's tortious conduct was a necessary element of a set of actual antecedent conditions that did not include the subsequent condition (the second bridge or the second injury) and was sufficient for the occurrence of the economic loss. The existence of the subsequent condition did not prevent this set from being sufficient to produce the economic loss. On the other hand, neither the second bridge nor the second injury was a necessary element of any set of *actual* antecedent conditions that was sufficient for the occurrence of the loss. The second bridge will cause a delay only if the boat reaches it, but the boat did not reach it.²⁵⁹ The second injury will cause total disability and resulting economic loss only if the victim is not already totally disabled,

254. *Id.* at 237-39.

255. *Id.* at 124, 206-07, 239-45.

256. H.L.A. HART & A. HONORÉ, *supra* note 10, at 118, 190, 224-25.

257. H.L.A. HART & T. HONORÉ, *supra* note 8, at 245.

258. *Id.* at 246-48, 250-51.

259. See L. GREEN, *supra* note 2, at 168-70.

but this also was not one of the actual conditions. In each instance, the defendant's tortious act is a preemptive cause of the injury and the resulting economic loss, and the subsequent preempted condition is not a cause of the economic loss.²⁶⁰

4. *Distinguishing the Damages Issue: The Successive-Injury and Overwhelming-Force Cases*

The successive-injury cases have engendered much debate and confusion, particularly in the Commonwealth countries, where the legal community seems unable to free itself from the but-for concept of causation.²⁶¹ The causal situation is clear in these cases. The first injury caused the economic loss; the second did not. The issue is not causal. It is a proximate-cause issue of policy or principle that is most appropriately placed under the heading of damages, and it also arises in the duplicative-causation cases. The issue is whether a defendant who has tortiously caused injury to the plaintiff nevertheless should be absolved from liability if the injury would have occurred anyway as a result of independent duplicative or preempted conditions.

Courts generally absolve the defendant from liability if he proves that the injury would have occurred anyway as a result of independent nontortious conditions.²⁶² In such a case, the plaintiff's corrective-justice claim—that he would not have been injured if not for the tortious conduct of others—fails. On the other hand, if the duplicative or preempted conditions also resulted from tortious conduct, the plaintiff's corrective-justice claim is intact.²⁶³ Nevertheless, it may seem unfair to hold a defendant liable for the entire injury if his contribution to the injury was relatively minor compared to the other contributing conditions. But this proximate-cause problem has been greatly lessened by the recent widespread liberalization of the rules governing contribution among defendants.²⁶⁴ Therefore, once the plaintiff has established that the tortious aspect of a certain defendant's conduct contributed to the injury, many

260. In the successive-injury case, the second condition is a cause of the second injury, and of any additional pain or costs resulting from the second injury.

261. *E.g.*, Jobling v. Associated Dairies, 1982 A.C. 794 (1981); J. FLEMING, *supra* note 4, at 174-76; Fraser & Howarth, *supra* note 227, at 136-37, 145-56; McGregor, *Successive Causes of Personal Injury*, 33 MOD. L. REV. 378 (1970); Strachan, *supra* note 7; Williams, *supra* note 21, at 75-79. For similar confusion by American writers, see Henderson, *supra* note 105, at 201-12; Peaslee, *supra* note 175.

262. *E.g.*, Kingston v. Chicago & N.W. Ry., 191 Wis. 610, 211 N.W. 913 (1927) (dicta); Cook v. Minneapolis, St. P. & S. Ste. M. Ry., 98 Wis. 624, 74 N.W. 561 (1898); Jobling v. Associated Dairies, 1982 A.C. 794 (1981). *But see RESTATEMENT (SECOND) OF TORTS* § 432(2) comment d (1965); Carpenter, *Concurrent Causation*, 83 U. PA. L. REV. 941 (1935).

263. *E.g.*, Kingston v. Chicago & N.W. Ry., 191 Wis. 610, 211 N.W. 913 (1927); Baker v. Willoughby, 1970 A.C. 467 (1969).

264. *E.g.*, American Motorcycle Ass'n v. Superior Court, 20 Cal. 3d 578, 578 P.2d 899, 146 Cal. Rptr. 182 (1978); Dole v. Dow Chem. Co., 30 N.Y.2d 143, 282 N.E.2d 288, 331 N.Y.S.2d 382

courts shift the burden to the defendant to establish that (1) the injury would have occurred anyway as a result of independent nontortious conditions, (2) he contributed to only a part of the injury, or (3) he is entitled to contribution from the other defendants based on relative tortious contribution.²⁶⁵

The causal issue is almost always confused with the policy-laden damages issue in the overwhelming-force cases involving passive conditions. When an overwhelming force combines with an *active* condition, the courts are less likely to confuse the two issues. Thus, in the merged-fires cases, the courts generally recognize that the defendant's smaller fire that combined with a much larger fire nevertheless contributed to the injury. If the defendant avoids liability, it is not due to a lack of causal contribution but rather due to noncausal limitations. For example, the defendant usually can avoid liability if he proves that the other fire was independently sufficient and of nontortious origin, or perhaps even if it was also of tortious origin if it was so overwhelming as to make the defendant's contribution relatively insignificant.²⁶⁶ On the other hand, the defendant is unlikely to escape liability even though his contribution was a very small part of the total cause if it was necessary for the consequence or, although unnecessary, was combined with a large number of similarly small contributions, as in the pollution cases.²⁶⁷

When the overwhelming force combines with a *passive* condition, however, most courts and commentators tend to fall back on the but-for test and erroneously deny causal contribution.²⁶⁸ For example, the courts deny that the defendant's negligent failure to construct an adequate dam or to keep runoff pipes free of debris was a cause of flooding that would have occurred given a normal storm, if the actual storm was an extraordinary one that would have caused flooding even if the defend-

(1972); Landes & Posner, *Joint and Multiple Tortfeasors: An Economic Analysis*, 9 J. LEGAL STUD. 517, 550-51 (1980); Robinson, *supra* note 7, at 716 & n.12.

265. In addition to the cases cited *supra* notes 262-64, see *Michie v. Great Lakes Steel Div.*, 495 F.2d 213 (6th Cir.), *cert. denied*, 419 U.S. 997 (1974); *Maddux v. Donaldson*, 362 Mich. 425, 108 N.W.2d 33 (1961). Some courts have adopted a "comparative causation" allocation formula in the strict product liability cases, since there is no negligence or fault as is required under the comparative-negligence (fault) rubric. *E.g.*, *General Motors Corp. v. Hopkins*, 548 S.W.2d 344 (Tex. 1977). However, from a purely causal standpoint, each necessary (or NESS) condition is equally a cause. At best, the comparative-causation formula is a camouflage for a "comparative responsibility" approach (surely a better and more helpful term), in which the types and degrees of tortious conduct involved and the policies underlying each type of tortious conduct are weighed to allocate liability. *E.g.*, H.L.A. HART & T. HONORÉ, *supra* note 8, at 232-34; see Robinson, *supra* note 7, at 758-68. At worst, the comparative-causation formula invites resort to mechanistic, generally inapplicable calculations of relative energy or force, which have little to do with the basic principles and goals of tort liability. *E.g.*, Epstein, *Defenses*, *supra* note 49, at 179-80; Epstein, *Reply*, *supra* note 49, at 478 n.5.

266. *E.g.*, *Kingston v. Chicago & N.W. Ry.*, 191 Wis. 610, 615-16, 211 N.W. 913, 915 (1927).

267. See cases cited *supra* note 239.

268. This point is noted in Malone, *supra* note 6, at 92-94.

ant had not been negligent.²⁶⁹ Similarly, they have stated that the defendant's negligent failure to remove a rotten pole is not a cause of injury when the pole is knocked down by a car that hits it with such force that it would have been knocked down even if it were not rotten.²⁷⁰

These are overdetermined-causation cases that are entirely analogous to the merged-fires cases and the weakened-cable case.²⁷¹ Thus, if the rotten pole broke at its weak point, the weak condition clearly contributed to the injury. It was a necessary element of a sufficient set of actual antecedent conditions that included an impact with at least enough force to knock down a rotten pole. The sufficiency of the set was not affected by the fact that the impact had a greater force, unless the impact knocked the pole over in its entirety rather than breaking it at its weak point. Similarly, the failure to remove debris from a dam's runoff pipes or to build it high enough to prevent flooding during a normal storm is a duplicative cause of flooding that occurs during an extraordinary storm. The unremoved debris or inadequate height is a necessary element in a sufficient set of actual antecedent conditions that includes an at least normal storm, and the sufficiency of this set is not affected by the fact that the storm was larger than normal.

A few courts have recognized that the defendant's tortious conduct contributed to the injury in these passive-condition cases. Indeed, some courts have held the defendant liable even when the injury would have occurred anyway as a result of independently sufficient nontortious conditions.²⁷² However, if the noncausal policy limitations adopted in the merged-fires and successive-injury cases are followed, the defendant should be able to avoid liability in such circumstances. These policy limitations, rather than the false denial of causation, explain the ultimate result in those cases in which liability is denied because the injury would have occurred anyway due to the extraordinary force of a storm.

Conversely, given the usual policy limitations, the defendant should not escape liability when the duplicative or preempted condition was also of tortious origin—for example, when the defendant's rotten pole is knocked down by a negligent driver and an innocent third party is injured. This case is the same as a merged-fires case in which the defend-

269. *E.g.*, *City of Piqua v. Morris*, 98 Ohio St. 42, 120 N.E. 300 (1918); RESTATEMENT (SECOND) OF TORTS § 432(1) illustration 2 (1965); A. BECHT & F. MILLER, *supra* note 8, at 90-93; L. GREEN, *supra* note 2, at 150-51; PROSSER & KEETON, *supra* note 7, § 41, at 265-66 & n.16. *Contra* RESTATEMENT (SECOND) OF TORTS § 450; Malone, *supra* note 6, at 92-94.

270. *E.g.*, *Gibson v. Garcia*, 96 Cal. App. 2d 681, 687, 216 P.2d 119, 123 (1950).

271. See *supra* text following notes 245 & 250.

272. *O'Connor v. Chicago, M. & St. P. Ry.*, 163 Wis. 653, 654, 158 N.W. 343, 344 (1916) (weak tree blown down by storm that would have blown down sound tree); *Nitro-Phosphate & Odam's Chem. Manure v. London & St. Katherine Docks*, 9 Ch. D. 503, 527 (C.A. 1878) (levee built too low overflowed by extraordinarily high tide that would have overflowed levee built to proper height); see *Carpenter, supra* note 127, at 418-19.

ant's insufficient fire joins with an independently sufficient fire, also of tortious origin. In each case, both conditions are tortious duplicative causes of the injury, and liability should be imposed jointly on both tortfeasors with a right of contribution.²⁷³

5. *Theft, Nonuse, or Misuse of Defective or Missing Safety Devices*

Some of the most difficult overdetermined-causation cases are the preemptive-causation cases involving theft, nonuse, or misuse of defective or missing safety devices. A familiar example is the case in which *C* negligently failed to discover and repair defective brakes in a car that he rented to *D*, and *D* negligently failed to try to use the brakes to avoid running into *P*. It is assumed that the injury to *P* would have been avoided if and only if *C* had repaired the brakes and *D* had tried to use them. As we saw above, Green asserts that *C*'s conduct as a whole and *D*'s conduct as a whole were each causes of the injury, but that only *C* should be held liable.²⁷⁴ Becht, Miller, Prosser, and Fleming assert that neither *C*'s negligence nor *D*'s negligence was a cause of the injury, but that both *C* and *D* should be held liable.²⁷⁵

Under the NESS test, it is clear that *D*'s negligence was a preemptive cause of *P*'s injury, and that *C*'s negligence did not contribute to the injury. *D*'s failure to try to use the brakes was necessary for the sufficiency of a set of actual antecedent conditions that did not include *C*'s failure to repair the brakes, and the sufficiency of this set was not affected by *C*'s failure to repair the brakes. A failure to try to use brakes will have a negative causal effect whether or not the brakes are defective. On the other hand, *C*'s failure to repair the brakes was not a necessary element of any set of antecedent *actual* conditions that was sufficient for the occurrence of the injury. Defective brakes will have an actual causal effect only if someone tries to use them, but that was not an actual condition here. The potential negative causal effect of *C*'s failure to repair the brakes was preempted by *D*'s failure to try to use them.²⁷⁶

The same analysis applies in every case involving nonuse or misuse of a missing or defective safety device, unless the actor did not try to use the device because he knew it was missing or defective.²⁷⁷ For example, failure to provide a fire escape was not a cause of *P*'s injury if *P* could not have gotten to the fire escape anyway or would not have tried to get to it

273. See *supra* text accompanying notes 262-67.

274. See *supra* text accompanying notes 108-15.

275. See *supra* note 223 and accompanying text.

276. E.g., *Saunders Sys. Birmingham Co. v. Adams*, 217 Ala. 621, 117 So. 72 (1928).

277. In the latter situation, the causal effect of the missing or defective device is a mental rather than a physical one, which actually exists and is not preempted but rather completed by the actor's failure to try to use the device. See *supra* notes 117 & 212.

even if it existed.²⁷⁸ Similarly, failure to provide safety equipment to a worker is not a cause of his injury if he would not have used it anyway,²⁷⁹ and failure to pack a parachute properly is not a cause of death if the deceased never pulled the ripcord to release the parachute.

A somewhat different situation is presented by McLaughlin's hypothetical in which *D* empties *P*'s water keg and refills it with salt before *P* is dropped into the middle of a desert. *C* subsequently steals the keg, thinking that it contains water, and *P* dies of thirst.²⁸⁰ This case is different from the previous cases because it is assumed that *P* would have used the water ("safety device") if it had been available. *D*'s emptying the keg made the water unavailable and was a preemptive cause of *P*'s death. *C*'s theft of the keg was not a cause of *P*'s death. Its potential effect (making the water unavailable) depended on the keg's having water in it when it was stolen, but that was not an actual condition.²⁸¹

Hart and Honoré modify the facts by having *D* poison the water in the keg rather than replacing it with salt. They conclude that neither *C* nor *D* was a cause of *P*'s death, but only because they equate causing death with shortening life and note that *P* would have died sooner by poisoning if *C* had not stolen the keg.²⁸² As Becht and Miller point out, and as a NESS analysis demonstrates, the lack of water (fluid) caused the death, and it is irrelevant that *P* would have died sooner if he had drunk the poisoned water.²⁸³ He in fact did not drink the poisoned water, because *C* stole it. *C*'s theft caused *P*'s death and preempted the potential causal effect of *D*'s poisoning the water.

It should be clear by now that the NESS test not only resolves but also clarifies and illuminates the causal issues in the problematic causation cases that have plagued tort scholars for generations. It does so because it is not just a test for causation, but is itself the meaning of causation. When there is no overdetermined-causation problem—that is, when there is only one actual or hypothetical sufficient set of conditions for a particular event—the NESS test collapses into the simple, tradi-

278. *E.g.*, *Weeks v. McNulty*, 101 Tenn. 495, 48 S.W. 809 (1898). *Contra* *Fraser & Howarth*, *supra* note 227, at 135-36.

279. *E.g.*, *M'Williams v. Sir Williams Arrol Ltd.*, 1962 Sess. Cas. 70 (H.L.); *Qualcast (Wolverhampton) Ltd. v. Haynes*, 1959 A.C. 743; *H.L.A. HART & T. HONORÉ*, *supra* note 8, at 127. *Contra* *Fraser & Howarth*, *supra* note 227, at 135-36.

280. *McLaughlin*, *supra* note 175, at 155 n.25.

281. *See supra* text accompanying notes 250, 258-60. Becht and Miller conclude that neither *C* nor *D* was a cause. *A. BECHT & F. MILLER*, *supra* note 8, at 210. This erroneous conclusion results from their application of the but-for test to two factors that were both in the same causal sequence. *See supra* text accompanying notes 222-24.

282. *H.L.A. HART & T. HONORÉ*, *supra* note 8, at 239-40 & n.70.

283. *A. BECHT & F. MILLER*, *supra* note 8, at 205-10. Becht and Miller rely on the tautological "would not have died when and as he did" argument. *Id.* at 210; *see supra* text accompanying notes 176-78. The applicable NESS analysis is discussed *supra* text accompanying notes 249-50.

tional but-for test. But the concept of causation is much more subtle and complex than the but-for test. Thus, whenever the but-for test is not satisfied, the factfinder must utilize this more complex notion of causation, with its often subtle distinction between actual sufficiency and mere apparent (preempted) sufficiency, before the causal inquiry can be concluded.

F. The Factual Nature of the Causal Inquiry

Now that the content of the causal inquiry has been identified, we are in a better position to evaluate the frequent claims that the causal inquiry is dependent on policy considerations. In the first portion of this section, I argue, contrary to the commonly held view, that the hypothetical inquiry associated with the necessary-condition criterion does not introduce policy issues into the causal inquiry. In the second portion, I examine and reject the various arguments made by Wex Malone, who generally is thought to have proven that the causal inquiry is pervaded by policy considerations. I conclude that the causal inquiry is a factual, empirical inquiry that can be—and almost always is—kept distinct from the policy issues in tort adjudication.

1. The Hypothetical-Inquiry Argument

Many legal writers assert that the counterfactual analysis of necessity "takes the eye off the ball" by asking what would have happened, rather than focusing on what actually happened, and thereby injects policy issues into the causal inquiry. Some writers deplore this alleged effect of necessity analysis,²⁸⁴ while others embrace it.²⁸⁵

The message of the preceding sections of this Part, however, is that the notion of necessity is fundamental to the concept of causation, although it is subsidiary to the notion of sufficiency rather than being the direct and exclusive criterion of causal contribution as in the but-for test. Efforts to escape from necessary-condition analysis by applying the causal inquiry to the actor's conduct as a whole or resorting to unelaborated formulas or paradigm cases are therefore ultimately futile.²⁸⁶

The causal question is not simply "What happened?"²⁸⁷ but rather "How did it happen?" Since Hume, it has generally been accepted that there is no inherent causal force or quality in objects which can be directly observed. We observe only certain successions of events and

284. E.g., Green, *supra* note 158, at 605; Green, *supra* note 8, at 556-57, 559; Thode, *supra* note 100, at 426-27 & n.15, 431; Note, *supra* note 27, at 1531-33; see Epstein, *supra* note 13, at 160-61.

285. E.g., Malone, *supra* note 6, at 67-68 & n.9; sources cited *supra* note 7.

286. See *supra* text accompanying notes 49-99, 107-17 & 190-225.

287. E.g., W. PROSSER, *supra* note 1, § 41, at 237; Epstein, *supra* note 13, at 160; Morris, *supra* note 3, at 1088-89; Thode, *supra* note 100, at 431, 433-34.

conditions, and we infer a causal relation in a particular succession if we believe it instantiates an accepted causal law or generalization. Causal generalizations, in turn, incorporate the notion of necessary conditions.²⁸⁸ Thus, as legal and nonlegal philosophers have noted, the necessary-condition criterion and its implicit counterfactual assertions are part of the very meaning of causation.²⁸⁹

In tort adjudication, the causal inquiry is focused on the tortious aspect of the actor's conduct. The trier of fact must determine whether all the conditions that were necessary to make the defendant's conduct tortious, including the relevant circumstances as well as the particular act or omission, contributed to the injury.²⁹⁰ The counterfactual analysis of necessity serves two purposes in this causal inquiry. First, it reduces the chance that a tortious condition erroneously will be treated as part of a causal generalization that is believed to be applicable to the particular injury. Second, it reduces the chance that preemptive causes will be overlooked. Thus, the counterfactual analysis of necessity does not take the eye off the ball. It rather helps to ensure that the right ball is being used. The objection to counterfactual analysis of necessity per se is therefore misplaced.

Most objections to the counterfactual analysis of necessity have focused on the clearly erroneous denials of causation that would be compelled by the but-for test if it were applied to the overdetermined-causation cases.²⁹¹ This problem arises, however, only because the but-for test makes the necessary-condition criterion the exclusive criterion for causal contribution. The problem is eliminated under the NESS test, which subordinates the necessary-condition criterion to the sufficiency criterion.

However, some writers have insisted that the counterfactual analysis of necessity inherently depends on value judgments. Ironically, most of these writers believe that causal analysis is a factual analysis. They therefore reject the counterfactual analysis of necessity because they think that it introduces policy considerations into the causal analysis.²⁹² They do not realize that their arguments, if successful, would negate their claims

288. See *supra* text accompanying notes 228-30.

289. T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 119-20; H.L.A. HART & T. HONORÉ, *supra* note 8, at 14-15, 109-12; Mackie, *supra* note 21, at 30-31.

290. See *supra* text accompanying notes 135-45.

291. E.g., EPSTEIN, TORTS, *supra* note 4, at 271-72; Calabresi, *supra* note 9, at 86-87, 105-08; Green, *supra* note 158, at 604-05; Green, *supra* note 8, at 556-60; Landes & Posner, *supra* note 9, at 109-11, 134; Malone, *supra* note 6, at 67, 88-97; Rosenberg, *supra* note 7, at 863-64; Thode, *supra* note 100, at 431-33; Williams, *supra* note 21, at 75-79.

292. Cole, *supra* note 7, at 482-87, 505-06 & n.133, 768-74 & n.16; Epstein, *supra* note 13, at 160-61, 165-66; Thode, *supra* note 100, at 424-27, 431. Compare Malone, *supra* note 6, at 67-69 & n.9 (welcoming policy considerations into causal inquiry), with *id.* at 71 & n.25 (causal issue usually clear on the facts) and Malone, *supra* note 27, at 370-73 (causal issue always factual).

that the causal inquiry is factual, since the counterfactual analysis of necessity is an indispensable element of causal analysis.

The argument against counterfactual analysis is developed most fully by Robert Cole. Initially, he asserts that the very decision to employ one form of necessary-condition analysis rather than another is a policy decision. He claims that there is no independent, policy-neutral basis for choosing "necessary element of a sufficient set" (the NESS test), "necessary on the particular occasion" (the but-for test), "necessary on every occasion," or any other necessity formula.²⁹³ But, as we have seen,²⁹⁴ the "necessary element of a sufficient set" formula is the essence of the concept of causation. The other formulas do not capture the notion of causal contribution, although one of them (the but-for test) is an adequate proxy in most cases.

Cole also argues that the courts introduce policy considerations when they assume, as part of the counterfactual situation, that the defendant would have behaved reasonably. In particular, he argues that when the courts analyze what would have happened if the defendant had not been driving at eighty miles per hour, there is no neutral basis for them to pick the legal speed limit (or some lower reasonable speed if the legal speed was unsafe in the particular circumstances) as the basis for the counterfactual inquiry. They should instead consider each possible alternative speed that the defendant might have chosen other than eighty miles per hour and the changes in behavior (such as greater attentiveness or deference to other drivers) that might accompany each alternative speed. In order to decide whether the defendant's eighty miles per hour speed was a necessary condition for the accident, the court must consider all these alternatives in as much detail as possible. It must then determine whether some speed other than eighty miles per hour was not only noticeably more likely than all the other possible alternative speeds but also more likely to result in the accident. If such an alternative speed exists, the eighty miles per hour speed was not a necessary condition for the accident.²⁹⁵

Cole then notes that the "pyramiding of conjectural possibilities" inherent in any such detailed consideration of what might have happened if the defendant had not been driving at eighty miles per hour makes it impossible to reach any definite conclusion.²⁹⁶ When the courts select the legal speed as the counterfactual speed and do not consider behavioral changes that might be expected from someone driving at the legal

293. Cole, *supra* note 7, at 769-71 & nn.15 & 16.

294. See *supra* text accompanying notes 228-32.

295. Cole, *supra* note 7, at 769-70, 773-76, 783-84; see also Thode, *supra* note 100, at 426-27 & n.17.

296. Cole, *supra* note 7, at 776-77.

speed rather than at eighty miles per hour, they are using policy considerations to drastically reduce the range of conjecture about what would have happened if the defendant had not been driving at eighty miles per hour.²⁹⁷

This argument depends on the common but fallacious assumption that the courts are trying to determine what would or might have happened if the defendant had not been driving at eighty miles per hour. Insofar as the causal inquiry is concerned, the courts are not interested, and should not be interested, in that broad inquiry. They are interested only in determining the causal effect of the tortious aspect of the defendant's conduct. For example, in the speeding cases the courts want to know whether the excess speed (the tortious aspect) was a necessary element in some set of actual antecedent conditions that was sufficient for the occurrence of the accident. They therefore change *only* the tortious condition when setting up the counterfactual situation. They remove the speed in excess of the legal speed from the set of actual antecedent conditions to determine whether the set would have been sufficient without it. They want to know whether the tortious aspect (the excess speed) actually contributed to the accident. They do not want to know, as part of the causal inquiry, whether the defendant would have driven at the legal speed if he had not driven at eighty miles per hour, or whether he might have been more attentive or otherwise more careful if he had been driving at the legal speed. The causal issue is restricted to the narrow question of the actual effect of the excess speed given the other conditions (attentiveness, etc.) that actually existed.

Similarly, the courts do not change the actual conduct of the plaintiff or third parties when they set up the counterfactual situation. They are not using policy considerations, but rather careful causal methodology, when they remove only the defendant's excess speed in order to determine the causal effect of that excess speed in the actual circumstances.²⁹⁸

Of course, once a court sets up the counterfactual situation by eliminating the particular tortious condition, it must then determine the hypothetical consequences of that change. As even Malone admits,²⁹⁹ in most cases there is little difficulty, for example, when the change is removing the explosive character of a substance or the act of firing a gun. The analysis becomes more complicated when human reactions to the changed situation must be estimated. Again, however, the analysis is usually fairly easy and not too speculative. People's reactions generally will be fairly predictable using causal generalizations in which there is a

297. *Id.* at 771, 784-85, 792-97; see also Malone, *supra* note 6, at 67.

298. *Contra* Thode, *supra* note 100, at 426 n.15.

299. Malone, *supra* note 6, at 71 & n.25.

high degree of confidence. Indeed, the defendant normally will stipulate that he would have reacted in the usual manner—for example, by applying his brakes if he would have had time to do so at the legal speed after seeing the plaintiff in his path. As Cole himself notes, it does not help the defendant's case to argue that one specific act of negligence did not contribute to the injury because he would have caused it anyway by some other negligent act or omission.³⁰⁰

In any event, the hypothetical inquiry is not policy-driven. Instead, empirically derived causal generalizations are used to estimate the most likely consequences of the counterfactual situation. Cole seems to believe that causal generalizations are not used in the counterfactual inquiry, or that the hypothetical development of the counterfactual situation must be specified in such detail as to make its occurrence (in that precise manner) improbable.³⁰¹ As philosophers have noted, however, counterfactual assertions are nothing more than "telescoped arguments" based on more or less well-grounded causal generalizations. The credibility of a particular counterfactual assertion depends on the extent of empirical support for the causal generalizations themselves and the extent to which the conditions that make up the counterfactual setting correspond to those that are specified in the causal generalizations.³⁰² These conditions, as Hart and Honoré have argued, must be specified at a fairly high level of generality if they are to serve as a broadly applicable causal generalization, rather than as a detailed unique description of a particular event with no causal explanatory force.³⁰³

2. *Malone's Arguments*

Malone attacks the factual nature of the causal inquiry differently. He does not challenge the process by which the counterfactual setting is established, and he recognizes that causal generalizations provide the foundation for any particular causal judgment. Instead, he presents an array of arguments to demonstrate that the policies underlying particular legal rules affect both causal judgments themselves and also the courts' selection of the degree of confidence that those judgments must attain in order to sustain liability in a specific case.

Malone notes that causal judgments, like any other evaluation of raw evidentiary data, involve interpretation and the drawing of inferences based on the past experience and "judging personality" of the per-

300. Cole, *supra* note 7, at 778-79; accord A. BECHT & F. MILLER, *supra* note 8, at 175-76.

301. Cole, *supra* note 7, at 771-74, 777, 784-91.

302. T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 119-20, 131-37, 145-49; Mackie, *supra* note 21, at 24-25, 30-31.

303. H.L.A. HART & T. HONORÉ, *supra* note 8, at 45-48; see T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 141; *infra* text accompanying notes 366-72. *Contra* Cole, *supra* note 7, at 791-92.

son making the judgment. He asserts that, particularly with respect to the causation issue, "the evaluation which [a person] will make of the new fact data will necessarily be affected by the purpose he is seeking to serve."³⁰⁴

Malone's principal argument in support of this assertion is the apparent effect of purpose on a person's selection of one of many contributing conditions as "the" cause. As seen in Part I, this argument confuses the tortious-conduct and proximate-cause inquiries with the causal inquiry. The argument demonstrates only that purpose affects the tortious-conduct and proximate-cause inquiries, not that it affects the causal inquiry.³⁰⁵

Malone also argues, however, that purpose enters into the causal inquiry itself. He observes that the process of forming and applying causal generalizations "is basically one of conjecture, for facts almost never represent themselves in identical patterns."³⁰⁶ Moreover, he notes, even constant repetition of an identical pattern is not *per se* sufficient to support a causal generalization. Something more is needed:

Previous observations by a stargazer may enable him to say with absolute assurance that when a given constellation rises in the heavens another constellation will soon come into view over the horizon. Yet, no one in this modern world would assume that the rising of the first constellation is a *cause* of the appearance of the second. The inevitable character of the sequence as gained from past observations is not enough to establish the relation of cause and effect. Other portions of our knowledge are too likely to get in the way of such a conclusion. There must be some acceptable point of affinity between the new event and old experience that is satisfactory, and this point of satisfaction can be established only in terms of the purpose toward which the whole process of decision is being directed.³⁰⁷

Indeed, something more than mere observation of repeated conjunction of the two events is needed. The something more is the belief, based on all our empirically derived knowledge, that the two events are linked through one or more causal generalizations as cause and consequence. Malone acknowledges that, regardless of the purpose of the inquiry, "*no one* in this modern world would assume that the rising of the first constellation is a *cause* of the appearance of the second. . . . Other portions of our *knowledge* are too likely to get in the way of such a conclusion." Although our judging and evaluating faculties are involved here, the inquiry is a factual one in which we judge or evaluate on the basis of empirically derived causal generalizations that are independent from and

304. Malone, *supra* note 6, at 61-62.

305. See *supra* text accompanying notes 17-27.

306. Malone, *supra* note 6, at 65.

307. *Id.* (emphasis in original).

not affected by the purpose of the particular causal inquiry.³⁰⁸

The second prong of Malone's attack focuses on the problematic causation cases. He claims that the courts have relaxed the actual-causation requirement in the overdetermined-causation cases and lowered or raised the burden of proof in the other types of problematic causation cases depending on their assessment of the policies underlying the rule that was violated.³⁰⁹

We have already seen that this claim fails with respect to the overdetermined-causation cases that commentators usually address.³¹⁰ Malone's own discussion of these cases persuasively demonstrates that the only "relaxation" of the actual-causation requirement that occurs is the abandonment of the inadequate but-for test in the face of clear evidence of causal contribution.³¹¹

Malone also discusses the "doubtful" causation cases, in which some safety device or precaution has been omitted that may well have contributed to the victim's injury, but the plaintiff is unable to establish that but for the defendant's omission the victim would not have been injured. Examples include cases in which a corpulent person slips and falls while hurrying down some unlit stairs, a child drowns in a pool without a lifeguard, a sailor falls overboard and drowns and the would-be rescuers are stymied by the lack of proper rescue equipment, or a pedestrian slips on an icy sidewalk with an irregular surface. Malone quotes language from several of these cases that suggests that the courts raised, lowered, or shifted the burden of proof on causation. He asserts that their decision to do so depended on the policies underlying the rule that was violated by the defendant.³¹²

But these cases are simply less familiar instances of the overdetermined-causation problem. A full reading of the cases in which the courts seem to lessen or shift the burden of proof regarding causation usually discloses that the court was explicitly or implicitly concerned that the defendant not escape liability under the governing but-for test when the jury reasonably could conclude from the evidence that it was more likely than not (the usual burden of proof in tort cases) that the defendant's omission of the safety precaution *contributed* to the injury.³¹³ Con-

308. Although there are undertones of epistemological skepticism in this portion of Malone's article, I do not believe that he is arguing that there are no real facts, or that we cannot perceive or agree on the existence of any facts because of our subjective filtering of sense data. See T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 16-17, 33-79, 88, 96-97, 112-15, 131-45, 152-55, 259-60; A. BECHT & F. MILLER, *supra* note 8, at 2-3, 9, 11.

309. Malone, *supra* note 6, at 67-72, 84-86, 88-89.

310. See *supra* text accompanying notes 235-60.

311. Malone, *supra* note 6, at 88-94; accord Carpenter, *supra* note 262, at 943, 952; see *supra* text accompanying notes 241-42.

312. Malone, *supra* note 6, at 68-79, 85-88, 94-96.

313. See cases cited *infra* note 335.

versely, the cases which seem to raise the burden of proof are generally either cases in which the court rigidly adheres to the but-for test despite good evidence of causal contribution, or cases in which the court believes the evidence is insufficient to support any well-grounded belief (rather than pure speculation) even on the issue of causal contribution.³¹⁴

For example, it usually is difficult to establish that proper lighting probably would have prevented the victim's falling down the stairs, or that proper construction or maintenance of a sidewalk to prevent irregularities in the surface probably would have prevented the victim from slipping on it when it was covered with ice. But the plaintiff often may be able to establish in such cases that the lack of light or the irregular surface more likely than not satisfied the NESS test and therefore contributed to the accident. Thus, the plaintiff may be able to prove that the lack of light was necessary for the sufficiency of a set of actual antecedent conditions which included a person of "at least *X* size" descending the stairs at a rate of "at least *Y* steps per minute," and that the sufficiency of this set was not preempted but rather reinforced by the existence of the additional conditions of greater size and faster descent. Similarly, the irregular surface might well have been necessary for the sufficiency of a set containing a sidewalk of "at least *X* slipperiness." As Leon Green clearly recognized, in these cases as well as in the more familiar over-determined-causation cases, the but-for test distracts attention from the issue of causal *contribution*.³¹⁵

As long as the but-for test is considered to be the applicable test of causation in these cases, there will be a strong urge to lessen or shift the burden of proof when the court perceives that causal contribution may well exist. While that approach was adopted in the original *Restatement of Torts*,³¹⁶ the *Restatement (Second)* states that the plaintiff bears the usual (more likely than not) burden of proof in these cases.³¹⁷ Although the *Restatement (Second)* deemphasizes but-for language in favor of substantial-factor language, the but-for language has not been entirely eliminated and the substantial-factor language is typically nebulous. Consequently, the overall discussion is hardly enlightening.³¹⁸

314. *E.g.*, *Taylor v. City of Yonkers*, 105 N.Y. 202, 209, 11 N.E. 642, 644 (1887) (rigid adherence to but-for test); *New York Cent. R.R. v. Grimstad*, 264 F. 334 (2d Cir. 1920) (pure speculation on causal contribution).

315. Green, *supra* note 8, at 556-59. *But see id.* at 560-61 (ratifying Malone's argument).

316. RESTatement OF TORTS § 432(1) comment c and accompanying illustrations (1934), cited in Malone, *supra* note 6, at 73 n.26.

317. RESTatement (SECOND) OF TORTS § 433B(1) comment b and accompanying illustrations (1965).

318. See A. BECHT & F. MILLER, *supra* note 8, at 135-37. The story in England is similar. For a brief period, the practice was to shift the burden to the defendant when a statutory duty had been breached to show that the precaution would not have averted the injury. *Vyner v. Waldenburg Bros.*, 1946 K.B. 50 (C.A. 1945). But then the House of Lords declared that the plaintiff must

The policies or principles being applied in these cases clearly are those underlying the actual causation requirement itself, rather than those underlying the particular rule that was violated. Malone's arguments to the contrary are not very persuasive. He argues, for example, that the shifting of the burden of proof in the hunting-accident cases, where it is known that one of the negligent defendants shot the plaintiff but not which one, is based on the very strict attitude toward firearm accidents, rather than the special causal proof problems in such alternative-causation cases.³¹⁹ He tries to rationalize defendants' greater success in the missing fire-escape cases than in the man-overboard cases by referring to the allegedly more restricted reach of the rule in the fire-escape cases. But his discussion demonstrates that the more restricted reach occurs not because of any difference in the strength or strictness of the policies, but because the evidence of actual causation in the fire-escape cases often shows that the failure to provide the fire escape did not contribute to the injury.³²⁰

Similarly, he asserts that the burden of proof is not lessened, but rather increased, in the medical-malpractice cases because of the important role fulfilled by doctors who apply "the tenets of an ever-changing and experimental science."³²¹ But again his discussion and, even more so, a note on which he relies³²² show that the burden of proof has not been raised in these cases. Instead, the special difficulties of proof in medical-malpractice cases, especially in the older cases,³²³ frequently have made it difficult for patients to establish actual causation. Expert testimony by other doctors normally is necessary, and their reluctance "to do more than express a conjecture on what would have happened with proper treatment constantly militates against the plaintiff's inter-

"prove his case by the ordinary standard of proof in civil actions: he must make it appear at least that on a balance of probabilities the breach of duty caused *or materially contributed to* his injury." Bonnington Castings Ltd. v. Wardlaw, 1956 A.C. 613, 620 (emphasis added) (opinion of Lord Reid). In subsequent cases the House of Lords has reaffirmed that contribution to the injury is sufficient to establish cause-in-fact, although the commentators continue to be mesmerized by the but-for test and to view the decisions as exceptions to the actual-causation requirement that were adopted for moral or policy reasons. *E.g.*, M'Ghee v. National Coal Bd., 1973 Sess. Cas. 37 (H.L. 1972), discussed in Fraser & Howarth, *supra* note 227, at 141-42; Weinrib, *supra* note 7. Weinrib takes a broader view of actual causation than most, but still argues that policy factors were determinative. For an earlier case very similar to *McGhee*, see Gardiner v. Motherwell Mach. & Scrap Co., 1961 Sess. Cas. 1 (H.L.). See *supra* text accompanying note 261.

319. Malone, *supra* note 6, at 82-85. The alternative-causation cases are discussed *infra* text accompanying notes 342-62.

320. Malone, *supra* note 6, at 78.

321. *Id.* at 86.

322. Annot., 13 A.L.R.2d 11, 21-24 (1950), cited in Malone, *supra* note 6, at 88 n.68.

323. Malone does not mention the conspiracy of silence among medical professionals. This practice was bolstered by the locality rule which limited testimony to doctors from the same locality and sometimes even the same school of practice. *See, e.g.*, Brune v. Belinkoff, 354 Mass. 102, 235 N.E.2d 793 (1968); Bowles v. Bourdon, 148 Tex. 1, 219 S.W.2d 779 (1949).

est."³²⁴ Usually they will be unable to state that the consequence would not have occurred but for the defendant doctor's negligent mistreatment or failure to treat, or even that the mistreatment or failure to treat contributed to the consequence. At best, they often may be able to testify only that the mistreatment or failure to treat reduced the plaintiff's chance of recovery—for example, from forty percent to twenty percent, or from eighty percent to seventy percent.³²⁵

The problems presented by these reduced-chance cases, as well as the similar increased-risk and alternative-cause (for example, hunting-accident) cases, are explored in Part III. For now, it is important to note that the problems are correctly perceived by courts and commentators as problems related to proof of causation, which are resolved by turning to the policies or principles that underlie the actual-causation requirement itself rather than the policies or principles that underlie the particular rule that was violated.³²⁶ Indeed, the few cases that have allowed recovery in the reduced-chance context generally have been medical-malpractice cases, despite Malone's contention that the "policy thrust" of the rules in this area is too "short and timid" to permit such recovery.³²⁷

Finally, it should be noted that shifting, lessening, or raising the burden of proof on causation does not make the causal inquiry itself any less factual. The causal inquiry is still an empirical, factual one that results in a certain degree of confidence that causation did or did not exist in a particular situation. The decision as to how high that degree of confidence must be, and who bears the burden of producing evidence to support that degree of confidence, is a distinct issue of policy or principle. In the criminal law, the prosecution bears the burden of establishing a very high degree of confidence: "beyond a reasonable doubt." In tort law, the plaintiff generally bears the burden of establishing a "more likely than not" degree of confidence. But in each context, the causal inquiry itself is a factual one.

Nevertheless, a demonstration of frequent ad hoc shifting or lessening of the burden of proof on causation would indicate that the actual-causation requirement, rather than being a fundamental criterion of tort liability, is just one more manipulable doctrine. But no such ad hoc

324. Malone, *supra* note 6, at 87; see Annot., *supra* note 322, at 22-23.

325. Malone, *supra* note 6, at 86-87.

326. See, e.g., sources cited *infra* notes 335, 339, 348; Delgado, *supra* note 7; Robinson, *supra* note 7.

327. Malone, *supra* note 6, at 81, 87. See, e.g., *Kallenberg v. Beth Israel Hosp.*, 45 A.D.2d 177, 357 N.Y.S.2d 508 (1974), *aff'd mem.*, 37 N.Y.2d 719, 337 N.E.2d 128, 374 N.Y.S.2d 615 (1975); *Herskovits v. Group Health Coop.*, 99 Wash. 2d 609, 664 P.2d 474 (1983). For useful reviews of the medical malpractice cases, see King, *Causation, Valuation, and Chance in Personal Injury Torts Involving Preexisting Conditions and Future Consequences*, 90 YALE L.J. 1353, 1365 n.38, 1367 n.45, 1368-69 n.53 (1981); Note, *Increased Risk of Harm: A New Standard for Sufficiency of Evidence of Causation in Medical Malpractice Cases*, 65 B.U.L. REV. 275, 281-92 (1985).

manipulation has been demonstrated. In almost all the cases that Malone discusses, the courts have been reacting to the inadequacies of the but-for test as a test of causal contribution. Very rarely have they allowed recovery when causal contribution to the injury could not be established under the more-likely-than-not standard.

The few exceptions—the reduced-chance, increased-risk and alternative-causation cases—are cases in which the courts are reluctant to allow a defendant to escape liability when his tortious conduct may well have contributed to the injury but it is impossible, given the nature of the situation, to produce evidence that he actually did contribute. In all these cases, as Malone admits, liability is never imposed if it is clear that the tortious aspect of the defendant's conduct did not contribute to the injury, no matter how strong the policy underlying the rule that was violated.³²⁸ I argue below, in Part III, that these cases are best understood as cases that define new legally protected interests, rather than as departures from the usual strictures of the actual-causation requirement.³²⁹

In sum, the causal inquiry is indeed a factual, policy-neutral inquiry, and the actual-causation requirement retains a secure position as a fundamental criterion of tort liability.³³⁰

III

CAUSATION, PROBABILITY, AND RISK: PROBLEMS ON THE FRONTIERS OF TORT LIABILITY

The most serious challenge to the actual-causation requirement has come from the increasing number of risk-exposure cases. These are cases in which one or more defendants tortiously exposed the plaintiff to the risk of an injury that subsequently occurred, but no proof exists that any particular defendant actually contributed to the injury. Many writers have asserted that the courts properly ignore or relax the actual-causation requirement in these risk-exposure cases in order to hold the defendants liable.³³¹ Recently, however, some writers have used the risk-exposure cases to support a much broader attack on the actual-causation requirement. Legal economists and others have argued that the traditional ex post inquiry into actual contribution to the injury should be replaced by an ex ante probabilistic assessment of increased risk in all

328. Malone, *supra* note 6, at 71, 77, 78; accord Carpenter, *supra* note 262, at 943, 947, 952.

329. See *infra* text accompanying notes 334-62.

330. Interestingly, Malone himself acknowledges these points in his later, less well-known writings, where he insists that the causal inquiry is "exclusively a fact inquiry" which "should be maintained utterly devoid of any policy overtones." Malone, *supra* note 27, at 371; see Note, *supra* note 27, at 1540-41 & nn.88 & 89.

331. See Calabresi, *supra* note 9; sources cited *supra* notes 6-8.

tort cases.³³²

I have demonstrated elsewhere that the probabilistic increased-risk concept cannot be substituted for the actual-causation requirement in the general run of cases without reaching results that are far removed from the traditional notions of liability that are applied by the courts.³³³ I argue here that, even in the risk-exposure cases, more satisfactory results are reached by adhering to the actual-causation requirement while recognizing a new type of injury—risk exposure—in certain narrowly circumscribed situations. The argument is developed by first analyzing the more traditional reduced-chance and increased-risk cases and then examining the currently more topical alternative-causation cases.

Finally, I attempt to clarify the fundamental distinction between ex ante statements of probability or increased risk and ex post statements of causal contribution. This distinction helps to clarify the risk-exposure cases and justify the courts' reluctance to accept naked statistical evidence as proof of causation.

A. *The Reduced-Chance and Increased-Risk Cases*

As was discussed above, although the plaintiff often cannot prove that the defendant's tortious conduct was a but-for cause of her injury, she frequently can prove that it more likely than not *contributed* to her injury.³³⁴ In such cases, the courts have held the defendant liable for the injury. Examples include cases in which the victim slips and falls on negligently unlit stairs, is washed overboard and there was a negligent failure to run lines around the boat to prevent such an occurrence, or is subjected to a negligent delay in treatment of an illness such as bronchial pneumonia which normally gets worse the longer it is untreated.³³⁵ Similarly, the courts have recognized that carcinogenic and toxic substances have cumulative, reinforcing effects, so that a defendant who exposes the victim to a substance which increases the risk of a certain cancer or illness is properly considered to have contributed to the subsequent occurrence of that cancer or illness, whether or not it might have occurred anyway.³³⁶

332. See Calabresi, *supra* note 9; Fraser & Howarth, *supra* note 227; Landes & Posner, *supra* note 9; Rizzo, *supra* note 227; see also G. CHRISTIE, *supra* note 4, at 246; Shavell, *supra* note 9.

333. Wright, *supra* note 9, at 452-55.

334. See *supra* text accompanying notes 312-27.

335. E.g., Zinnel v. United States Shipping Bd. Emergency Fleet Corp., 10 F.2d 47 (2d Cir. 1925) (missing lifeline); Reynolds v. Texas & Pac. Ry., 37 La. Ann. 694 (1885) (unlit stairway); Dunham v. Village of Canisteo, 303 N.Y. 498, 104 N.E.2d 872 (1952) (delay in treatment); RESTATEMENT (SECOND) OF TORTS § 433B(1) comment b and accompanying illustrations (1965); Green, *supra* note 8, at 558-59 & n.41.

336. E.g., McAllister v. Workmen's Compensation Appeals Bd., 69 Cal. 2d 408, 445 P.2d 313, 71 Cal. Rptr. 697 (1968); M'Ghee v. National Coal Bd., 1973 Sess. Cas. 37 (H.L. 1972); Clarkson v. Modern Foundries, [1957] 1 W.L.R. 1210 (Leeds Assizes); see Note, *Increased Risk of Cancer as an*

In other cases, however, the plaintiff may be unable to prove that the tortious aspect of the defendant's conduct more likely than not contributed to the injury. He may be able to prove only that the tortious aspect—for example, failure to provide the proper lifesaving equipment or to properly diagnose and treat the patient—reduced the victim's chance of avoiding the injury from, say, forty percent to twenty percent. In these cases, unlike the carcinogenic-exposure cases, there is no incremental, cumulative contribution. The omitted safeguard or treatment either would or would not have prevented the injury.

One approach to these cases, when the plaintiff cannot establish that the tortious aspect more likely than not contributed to the injury, is to reduce the plaintiff's burden of proof to permit recovery even though the degree of confidence in causal contribution is less than fifty-one percent. This approach, which has been adopted by a few courts,³³⁷ weakens the actual-causation requirement. It does not, however, make the causal inquiry any less factual.

A second approach is to allow recovery for the reduced chance of avoiding the injury, rather than for the injury itself. To apply this approach, the courts would have to recognize the chance of avoiding injury, *in a case where injury actually occurs*, as a legally protected interest. If the chance interest is deemed to be legally protected, at least against "significant impairments,"³³⁸ a court must then address the actual causation issue: did the tortious aspect of the defendant's conduct more likely than not contribute to a (significant) impairment of the chance interest in this case? If not, the court should not permit the plaintiff to recover, no matter how strong the policies underlying the rule that was violated.

The second approach preserves the integrity of the actual-causation requirement, while also addressing the basic issues in these cases. Should the defendant completely escape liability when his tortious conduct reduced the victim's chance of avoiding an injury that has actually occurred? This is an issue of policy or principle concerning the types of interests that should be legally protected. It should be faced and resolved explicitly, not covertly by diluting the actual-causation requirement.

Actionable Injury, 18 GA. L. REV. 563, 580-81 n.68 (1984). The author of the Note, however, erroneously assumes that the defendant's contribution was the *sole* cause if it was independently sufficient for the occurrence of the cancer. *Id.*

A defendant who has tortiously contributed to a certain instance of cancer or illness may seek contribution from other contributing tortfeasors, or from the plaintiff herself if the plaintiff also contributed (e.g., by smoking). The defendant may even escape liability entirely if he can prove that the cancer or illness would have occurred anyway as a result of independent nontortious contributing conditions. *See supra* text accompanying notes 258-73.

337. *See supra* note 327.

338. Compare the "significant harm" limitation applicable in private nuisance cases. RESTATEMENT (SECOND) OF TORTS § 821F (1979).

When the policy issues are carefully distinguished from the causal issue, the liability analysis is greatly improved. For example, if it is admitted that the interest being protected in these cases is the chance interest, recovery should be limited to the value of the lost chance rather than all the damages attributable to the injury itself. A practical method of valuing the lost chance is to multiply the total damages attributable to the injury by the amount of reduction in the chance of avoiding it.

The second approach appears to be gaining recognition as the preferable one in these cases.³³⁹ Under this approach, a plaintiff who is actually injured can recover for the injury itself if he can prove that the tortious aspect of the defendant's conduct contributed to the injury. If he cannot prove this, but can prove that the tortious aspect contributed to a specific reduction in his chance of avoiding the injury, he can recover for the reduction in the chance.

Some writers have argued that the courts should go even further. They would allow recovery for risk-imposition even when no tangible injury has yet resulted from the imposition of the risk.³⁴⁰ Such a dramatic extension of the category of protected interests would raise serious theoretical and practical problems.³⁴¹ In any event, the issues involved are not causal ones but questions of policy or principle concerning the types of interests that are to be protected against invasion. They will be resolved most clearly if they are seen as such.

B. The Alternative-Causation Cases

In the alternative-causation cases, several tortfeasors each independently exposed the victim to the risk of a specific injury that subsequently occurred. It is clear, however, that only one (or a few) of the tortfeasors actually contributed to the injury. The plaintiff is unable to establish which one (or few) actually contributed. For example, in the leading case of *Summers v. Tice*,³⁴² the plaintiff was struck in the eye by a single

339. See *O'Brien v. Stover*, 443 F.2d 1013, 1018-19 (8th Cir. 1971); *James v. United States*, 483 F. Supp. 581, 587 (N.D. Cal. 1980); *Herskovits v. Group Health Coop.*, 99 Wash. 2d 609, 622-36, 664 P.2d 474, 480-87 (1983) (Pearson, Stafford, and Utter, JJ., and Williams, C.J., concurring); PROSSER & KEETON, *supra* note 7, § 41, at 272; Green, *supra* note 8, at 558-59; King, *supra* note 327, at 1376-87; Malone, *supra* note 6, at 80-81.

340. Note, *supra* note 336; Note, *Increased Risk of Disease from Hazardous Waste: A Proposal for Judicial Relief*, 60 WASH. L. REV. 635 (1985).

341. Dworkin, *Fear of Disease and Delayed Manifestation of Injuries: A Solution or a Pandora's Box?*, 53 FORDHAM L. REV. 527 (1984). The courts have yet to accept this argument, even those courts with the most liberal attitudes toward liability. For a recent example, see *Collins v. Eli Lilly Co.*, 116 Wis. 2d 166, 191 n.10, 342 N.W.2d 37, 49 n.10, cert. denied, 105 S. Ct. 107 (1984), discussed *infra* text accompanying notes 356-57. It should be noted that tangible injury includes emotional distress, as in the "cancerphobia" cases. E.g., *Ferrara v. Galluchio*, 5 N.Y.2d 16, 152 N.E.2d 249 (1958).

342. 33 Cal. 2d 80, 199 P.2d 1 (1948).

pellet when the two defendants independently fired their shotguns in his direction. Obviously, only one of the defendants actually caused the plaintiff's injury, but there was no way to determine from which gun the pellet came. The court held the defendants jointly and severally liable for the injury. It shifted the burden of proof on causation to them, on the ground that it was better to have the loss fall on the two negligent defendants than on the innocent plaintiff.³⁴³

Summers represents a significant departure from the usual application of the actual-causation requirement. Prosser incorrectly states that the court "merely extended the rule as to the burden of proof on the issue of apportionment of damages."³⁴⁴ The rule that Prosser cites shifts the burden to the defendants to allocate the loss among themselves only after the plaintiff has established that each defendant contributed to the injury.³⁴⁵ In *Summers*, however, we know that one of the defendants is being held liable for an injury to which he did not contribute. This knowledge also distinguishes *Summers* from the reduced-chance cases.

As the number of defendants increases, the departure from the usual principles of tort liability becomes ever more apparent and difficult to justify. Thus, if there are N defendants, we know that $N - 1$ of them are

343. This rationale for joint and several liability was adopted long before *Summers* in *Tidal Oil Co. v. Pease*, 153 Okla. 137, 139-40, 5 P.2d 389, 390-91 (1931) (cattle drank from one or both of two streams poisoned by different defendants); *accord* *Bowman v. Redding & Co.*, 449 F.2d 956, 967-68 (D.C. Cir. 1971); *Abel v. Eli Lilly & Co.*, 418 Mich. 311, 325-29, 343 N.W.2d 164, 170-72, *cert. denied*, 105 S. Ct. 123 (1984); *Cook v. Lewis*, [1952] 1 D.L.R. 1 (Can. 1951); RESTATEMENT (SECOND) OF TORTS § 433B(3) (1965).

Although the *Summers* court noted that ordinarily defendants have better access to evidence on causation, this was not true in the particular case, and thus was not the basis for the court's decision. *Sindell v. Abbott Labs.*, 26 Cal. 3d 588, 601-03, 607 P.2d 924, 929-30, 163 Cal. Rptr. 132, 137-38, *cert. denied*, 449 U.S. 912 (1980).

Some writers have tried to justify the imposition of liability in the alternative-causation cases by arguing that the act of each defendant, in conjunction with the other's act, deprived the plaintiff of a chance of establishing who caused the injury. *E.g.* A. BECHT & F. MILLER, *supra* note 8, at 105 n.155; Weinrib, *supra* note 7, at 525-26. But this argument proves too much. It could be applied to any case in which there are alternative potential causes, including those attributable to the plaintiff, innocent third parties, or natural conditions or events. Acceptance of the argument means that the defendant is being held liable because he prevented the plaintiff from being consoled by the knowledge that the defendant was not a cause of his injury. The argument has little merit and distracts attention from the real issues.

For example, in *Haft v. Lone Palm Hotel*, 3 Cal. 3d 756, 478 P.2d 465, 91 Cal. Rptr. 745 (1970), the argument is entirely gratuitous. The absence of a lifeguard almost surely contributed to the drownings, and the court properly insisted that the duty to have a lifeguard remained in effect in the absence of the statutorily specified warning. The warning, if provided, would simply have reinstated the assumption-of-risk defense that ordinarily could not be invoked by someone who failed to comply with the statutory lifeguard requirement. *See Osborne v. Salvation Army*, 107 F.2d 929, 931-32 (2d Cir. 1939) (no assumption of risk if victim was member of class that statute was meant to protect); *Rovegno v. San Jose Knights of Columbus Hall Ass'n*, 108 Cal. App. 591, 291 P. 848 (1930) (absence of lifeguard contributed to drowning).

344. W. PROSSER, *supra* note 1, § 41, at 243 n.56.

345. *See* sources cited *supra* note 264.

being held liable even though they did not contribute to the injury. As N increases, the probability that any particular defendant is the one who caused the injury becomes increasingly small. The argument that the negligent defendants, rather than the innocent plaintiff, should pay for the injury becomes more a penal argument than a tort argument. As Professors Harper and James note, "the only fault which should be relevant to the question of civil liability is fault which has caused (or contributed to causing) the harm."³⁴⁶ The argument even runs counter to a basic tenet of the criminal law. In effect, the argument asserts that it is better that $N - 1$ defendants who did not contribute to the injury should pay for it than that the one who did cause it should go free. Moreover, the one who did cause it still may pay for little or none of the damages, due to contribution or insolvency.

The *Summers* rule disadvantages plaintiffs as well as defendants. The underlying rationale of *Summers* requires that all the alternative causes of the injury have been tortious and that all the tortfeasors be joined as defendants. These requirements ensure that the injury was tortiously caused and that the person who actually caused the injury is before the court. Thus, in a case in which forty persons were firing, but only seven of them were negligent, the *Summers* rationale was held inapplicable and the plaintiff obtained no compensation.³⁴⁷

These problems with the alternative-liability rationale of *Summers* have had to be confronted in the recent DES cases. In these cases, the plaintiff alleges that she contracted cancer as a result of her mother's ingesting the drug DES while pregnant with her. Typically, the plaintiff is unable to identify the particular firm or firms who supplied the DES that her mother ingested twenty or so years earlier. Because there are hundreds of firms that supplied DES, many of which no longer exist, most courts recognize that the *Summers* rationale and requirements break down in these cases.³⁴⁸

In *Sindell v. Abbott Laboratories*,³⁴⁹ the California Supreme Court modified the *Summers* approach to accommodate the most obvious procedural and fairness problems. It stated that the plaintiff need not identify

346. 2 F. HARPER & F. JAMES, *supra* note 2, § 20.2 n.24 comment, at 95 (Supp. 1968).

347. *Burton v. Waller*, 502 F.2d 1261 (5th Cir. 1974), cert. denied, 420 U.S. 964 (1975).

348. E.g., *Sindell v. Abbott Labs.*, 26 Cal. 3d 588, 603-04, 607 P.2d 924, 930-31, 163 Cal. Rptr. 132, 138-39, cert. denied, 449 U.S. 912 (1980); *Martin v. Abbott Labs.*, 102 Wash. 2d 581, 591-95, 689 P.2d 368, 375-77 (1984); *Collins v. Eli Lilly Co.*, 116 Wis. 2d 166, 183-84, 342 N.W.2d 37, 45-46, cert. denied, 105 S. Ct. 107 (1984). But see *Abel v. Eli Lilly & Co.*, 418 Mich. 311, 329-35, 343 N.W.2d 164, 172-75 (approving use of the *Summers* approach in the DES context), cert. denied, 105 S. Ct. 123 (1984). However, the court, referring to some of the procedural and fairness problems, reserved judgment on whether modifications should be made to the approach and on the "validity of any verdict that may result." *Abel*, 418 Mich. at 331 n.14, 339-40, 343 N.W.2d at 173 n.14, 177 (emphasis in original).

349. 26 Cal. 3d 588, 607 P.2d 924, 163 Cal. Rptr. 132, cert. denied, 449 U.S. 912 (1980).

and join all the firms that were operating in the relevant market at the time that her mother took the DES, but only enough firms to account for a "substantial percentage" (undefined but apparently between fifty and seventy-five percent) of that market. In addition, it replaced the defendants' joint and several liability, which holds each defendant responsible for the entire injury, with several (separate) liability based on each defendant's proportionate share of the market.

The substantial-percentage requirement apparently is intended to ensure that there is a substantial probability that the firm which actually caused the particular plaintiff's injury is before the court. The limitation of each defendant's liability to a proportion of the damages equivalent to its share of the market is intended to ensure that, over a large number of cases, the defendant is in effect being held liable for the injuries that it actually caused. That is, the two modifications of the *Summers* approach are intended to replicate the actual-causation requirement as applied to DES injuries in the aggregate, so that each individual plaintiff recovers for her injury and each defendant firm is held liable, in the aggregate, only for the injuries that it caused.³⁵⁰

However, two aspects of the *Sindell* approach undermine its ability to replicate the actual-causation requirement in the aggregate. First, the *Sindell* court apparently would allocate one hundred percent of the damages for each injury among the joined defendants based on their relative market shares, even though the joined defendants usually would represent less than one-hundred percent of the market.³⁵¹ Thus, the joined defendants (usually the largest firms) will be held liable for more than their expected share of the total DES damages. Second, unless the total number of DES injuries is very large, there will be only a very rough correlation between the injuries actually caused in the aggregate by a particular firm and those for which it is held liable under the probabilistic market-share approach. If there is only a relatively small number of injuries in the aggregate, the extent to which the actual distribution of injuries among firms varies from the predicted distribution, based on market share, may be quite substantial.³⁵²

The *Sindell* result is much more plausible theoretically if it is viewed as liability for having tortiously exposed the plaintiff to a certain type of risk, rather than liability for having caused a certain proportion of the actual injuries themselves. In effect, the *Sindell* court recognized a new

350. *Id.* at 613, 607 P.2d at 938, 163 Cal. Rptr. at 146; see Rosenberg, *supra* note 7, at 873, 881-83.

351. 26 Cal. 3d at 612-13, 617, 607 P.2d at 937, 940, 163 Cal. Rptr. at 145, 148.

352. For example, although there is an equal probability of a "heads" or "tails" on each toss of a coin, the fewer times the coin is tossed, the less likely it is that the actual result of the series of tosses will be approximately equal numbers of "heads" and "tails." See, e.g., H. BRUNK, AN INTRODUCTION TO MATHEMATICAL STATISTICS 153 (3d ed. 1975).

legal injury: tortious exposure to a risk that possibly led to a subsequent injury.³⁵³ A firm's market share of sales of an identical generic product such as DES precisely reflects its contribution to the aggregate risk to which the plaintiff and all others in the relevant market were tortiously exposed.³⁵⁴ If the plaintiff is injured as a result of her exposure to that aggregate risk, she has a cause of action against each firm which tortiously contributed to the risk.

As in the reduced-chance cases, the damages attributable to such risk-exposure can be calculated by multiplying the injury damages by the firm's proportionate contribution to the aggregate risk. Again, if the plaintiff can prove that a certain firm actually caused her injury by specifically identifying it as the source of the DES that her mother ingested, she should be able to recover her full damages from that firm. Conversely, if a certain firm can prove that it did not provide any of the DES that her mother ingested, it cannot be held liable for her injury or for having exposed her to a risk that possibly led to the injury.

Under this risk-contribution approach, it is no longer necessary to join enough defendants to account for a substantial percentage of the market, since the focus is on contribution to the risk rather than contribution to the DES injury. For the same reason, the approach avoids the dissenting judges' concern in *Sindell* that the defendants are being held liable without proof of actual causation.³⁵⁵ In each case, a defendant firm is held liable only if the plaintiff proves that the firm tortiously contributed to the aggregate risk that possibly led to the injury.

The risk-contribution approach to the DES cases has been adopted explicitly by the Wisconsin Supreme Court and implicitly by the Washington Supreme Court.³⁵⁶ Neither the Wisconsin court nor the Washington court requires that the joined defendants constitute a substantial percentage of the DES market. However, the Wisconsin court, like *Sindell*, holds the joined defendants liable for all the plaintiff's damages. It apparently allocates the liability severally among the joined defendants using comparative negligence principles as well as relative market shares.³⁵⁷ The Washington court initially holds the joined defendants severally liable for all the damages on a pro rata (equal share) basis.

353. Viewed in this way, the *Sindell* result is an extension of the risk-exposure approach to the reduced-chance cases. *See supra* text accompanying notes 338-39.

354. If there are risk-creating characteristics that vary among the products of different firms, they should be taken into account. *See Rosenberg, supra* note 7, at 867 & n.68.

355. *Sindell*, 26 Cal. 3d at 617, 607 P.2d at 940, 163 Cal. Rptr. at 148 (Richardson, Clark, and Manuel, JJ., dissenting).

356. *Martin v. Abbott Labs.*, 102 Wash. 2d 581, 594-95, 603-04, 689 P.2d 368, 377, 381-82 (1984); *Collins v. Eli Lilly Co.*, 116 Wis. 2d 166, 191 & n.10, 342 N.W.2d 37, 49 & n.10, *cert. denied*, 105 S. Ct. 107 (1984).

357. *Collins*, 116 Wis. 2d at 193-200, 342 N.W.2d at 50-53.

However, it allows each defendant that proves its absolute share of the total market to reduce its liability in accordance with that absolute market share. The remaining defendants split the remaining damages.³⁵⁸

One other group of alternative-causation cases remains to be discussed. The leading case is *Ybarra v. Spangard*,³⁵⁹ in which an unconscious patient's shoulder was injured during or immediately after an appendicitis operation. The California Supreme Court held that the patient could use the doctrine of res ipsa loquitur against all the doctors and nurses who had anything to do with him while he was unconscious, in order to raise an inference of negligent causation against the doctors and nurses as a group. Each would then be held jointly and severally liable if he or she failed to come forward with an absolving explanation.³⁶⁰ The court was clearly determined that a conspiracy of silence should not deprive the patient of his remedy.

Ybarra seems to go even further than *Summers*. It is not only unlikely that all the defendants in *Ybarra* contributed to the injury, but also unlikely that they all were negligent. However, the decision can be reconciled with the tortious-aspect causation requirement if it is viewed as imposing a form of enterprise liability. The court itself mentioned that all the defendants could be treated as permanent or temporary employees of the supervising surgeon or the hospital.³⁶¹ When all the defendants are connected through contractual or commercial relationships into a common enterprise and can adjust the risks and liabilities among themselves, and persons injured by that enterprise ordinarily will have a difficult time pinpointing the tortious source of the injury, it may be appropriate to treat the defendants as a group entity—an enterprise—which tortiously caused the injury, and to let the members of the enterprise allocate the liability among themselves or absolve themselves, as they see fit. The other situations to which the *Ybarra* rationale has been applied all fit this analysis.³⁶²

C. Distinguishing Probability Statements from Causal Attribution

The reduced-chance, increased-risk, and alternative-causation cases all raise questions about the connection that exists between various types of probability statements and the actual-causation requirement. These questions are part of an ongoing debate about the proper role of probabil-

358. *Martin*, 102 Wash. 2d at 604-06, 689 P.2d at 382-83.

359. 25 Cal. 2d 486, 154 P.2d 687 (1944).

360. *Id.* at 494, 154 P.2d at 691.

361. *Id.* at 492, 154 P.2d at 690. This seems to be the preferable approach, since the hospital and supervising surgeon can control and adjust the risks beforehand and obtain information afterward more easily than the nurses or orderlies can.

362. See cases cited in PROSSER & KEETON, *supra* note 7, § 39, at 253 nn. 22, 24 & 25.

istic data and arguments in legal adjudication in general.³⁶³

Due in large part to the lack of an adequate account of causation, the concept of causal explanation or attribution increasingly has been confused with predictive statements based either on causal generalizations or on mere statistical frequency. Some writers have asserted that the causation requirement itself should be interpreted in probabilistic terms. Under this notion of "causal linkage" or "probabilistic causation," the defendant's conduct is considered to have been a cause of the injury if it increased the risk of occurrence of the injury.³⁶⁴ Other writers have not gone so far, but have challenged the courts' reluctance to accept purely statistical data as evidence of causation or identification. These writers argue that all evidence, whether it is "particularistic" empirical data or pure statistics, is ultimately probabilistic, so that the courts' attitude that particularistic data is somehow qualitatively different, more reliable, or more trustworthy than statistical data is unsupportable.³⁶⁵

These arguments ignore some fundamental distinctions among causal explanations, causal predictions, and mere statistical reports or bets. They erroneously assume that because probability judgments underlie each of these types of statements, that is all that is meant by any of them. As will be seen, only causal explanations supply the attributive element that is essential for systems of liability based on individual responsibility. Mere statements of increased probability, even when based on causal generalizations, are by themselves insufficient to establish that attributive element. Probabilities based on causal generalizations, however, may help to distinguish suggested rival explanations. Pure sta-

363. E.g., L. COHEN, *THE PROBABLE AND THE PROVABLE* (1977); M. FINKELSTEIN, *QUANTITATIVE METHODS IN LAW* (1978); Jaffee, *Of Probativity and Probability: Statistics, Scientific Evidence, and the Calculus of Chance at Trial*, 46 U. PITTS. L. REV. 925 (1985); Kaye, *The Laws of Probability and the Law of the Land*, 47 U. CHI. L. REV. 34 (1979); Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L.J. 101; Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1377-90 (1985); Rosenberg, *supra* note 7, at 855-59, 869-74; Saks & Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 LAW & SOC'Y REV. 123 (1981); Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971); Williams, *The Mathematics of Proof* (pt. 1), 1979 CRIM. L. REV. 297; Brilmayer & Kornhauser, *Quantitative Methods and Legal Decisions* (Book Review), 46 U. CHI. L. REV. 116 (1978).

364. E.g., Calabresi, *supra* note 9, at 71-72; Fraser & Howarth, *supra* note 227, at 137-41; Landes & Posner, *supra* note 9, at 111-16, 134; Rizzo, *supra* note 178, at 1009-16, 1037-38; Shavell, *supra* note 9, at 468-69; see G. CHRISTIE, *supra* note 4, at 246; Robinson, *supra* note 7, at 758-60, 764-65. For my criticism of the views of Calabresi, Landes, Posner, and Shavell, see Wright, *supra* note 9.

365. E.g., Kaye, *The Laws of Probability and the Law of the Land*, 47 U. CHI. L. REV. 34 (1979); Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, 1982 AM. B. FOUND. RESEARCH J. 487, 488-89, 492 n.22, 514; Rosenberg, *supra* note 7, at 870-73; Saks & Kidd, *supra* note 363, at 149-54; Tribe, *supra* note 363, at 1344-50, 1361 n.102.

tistical data generally cannot perform even this function and therefore properly are rejected as evidence of causation or identification.

Causal explanations and causal predictions are applications of causal generalizations to empirical data. The causal generalizations themselves are derived inductively from empirical observations of repeated successions of conditions and events. A causal generalization, however, is not just a statement of observed statistical correlation. The essence of a *causal* generalization is the belief that we attach to the generalization: the belief in its causal or lawlike character. A causal generalization asserts that the antecedent conditions produce or cause the subsequent event—that they are necessary elements of a set of conditions that is sufficient for the occurrence of the event. It is precisely this quality of causal generalizations that distinguishes them from mere statistical reports and gives them explanatory and predictive power.³⁶⁶

When we seek to understand and explain the occurrence of an event, we compare the observed sequence of actual events (the "particularistic" empirical data) with the antecedent conditions and the associated consequence that are specified in the causal generalization. Usually, very few of the necessary antecedent conditions must be observed in addition to the consequence to induce us to believe that the causal generalization is applicable in the particular circumstances. This is true even when the ex ante probability of the consequence given these circumstances was very low—for example, *A*, a very poor marksman, shoots *B* from a great distance with a crooked gun in a high wind under conditions of poor visibility.³⁶⁷ The empirical observation of the consequence itself is a "vital part of the evidence which supports an explanation of that event."³⁶⁸

Indeed, as Hart and Honoré have stated, the force of a particular causal explanation usually depends on its referring to only a few of the antecedent conditions and its relying, explicitly or implicitly, on causal generalizations that are framed in broad or even platitudeous terms. It is unnecessary, even if it were possible, to explain a particular occurrence by detailing all the antecedent conditions so that we finally arrive at a universally true statement of invariable and unconditional causal connection. As the precision and detail of the description of all the antecedent conditions increases, our ability to predict the effect improves. Beyond a certain point, however, the explanatory force of the description does not improve, but rather lessens as it increasingly becomes a description of a unique event rather than an instance of some broad generalization appli-

366. T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 131-45.

367. See H.L.A. HART & T. HONORÉ, *supra* note 8, at 31-32, 44-48, 487.

368. Mackie, *supra* note 21, at 36-37.

cable to a multiplicity of different occurrences.³⁶⁹

On the other hand, when we use causal generalizations for predictive purposes, we obviously cannot rely on observation of the consequence that is being predicted. We must engage in a much more detailed investigation and explicit description of the actual antecedent conditions in order to have any substantial degree of confidence in our prediction.³⁷⁰ We are not able, as when explaining an event that has already occurred, to ignore all the possibilities concerning what might happen in the light of the existing conditions and possible future conditions.

It is important to distinguish causal prediction from causal explanation. When we are trying to explain a particular occurrence, the statement that *A* usually or almost always causes *B* is a predictive statement, which is not per se sufficient to persuade us that *A* caused *B* in this specific instance. After all, the "usually" or "almost always" qualifier admits that there are instances in which *A* does not cause *B*, and this occurrence may well be one of the exception cases. In order to be satisfied that the causal generalization "*A* causes *B*" is applicable in this specific instance, we must first distinguish any rival explanations that are suggested by the known or assumed (normal) circumstances.³⁷¹

It is at this point that the counterfactual analysis of necessity often enters into the causal inquiry. Usually the speculative element of the counterfactual analysis is minimal since only a single antecedent condition—one of the conditions that constituted the tortious aspect of the actor's conduct or activity—is changed in setting up the counterfactual situation, and the alternative explanations, if any, suggested by the counterfactual situation are generally few and have very different probabilities. However, when the changed condition generates a number of possible explanations that have similar probabilities, the speculative, predictive element becomes dominant and there will be little or no confidence in any particular explanation.³⁷²

Thus, in the doubtful-causation cases (man overboard, unlit steps, and so forth) the plaintiff establishes that the tortious condition was a cause of the victim's injury only if he can convince the trier of fact that his suggested causal explanation is more plausible than any suggested alternative causal explanations. If he is only able to prove that the tortious condition reduced the chance of avoiding the injury (increased the risk of its occurrence) but cannot sufficiently distinguish possible alterna-

369. H.L.A. HART & T. HONORÉ, *supra* note 8, at 44-47; see T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 141, 295-99, 301-02.

370. H.L.A. HART & T. HONORÉ, *supra* note 8, at 47; Mackie, *supra* note 21, at 37.

371. H.L.A. HART & T. HONORÉ, *supra* note 8, at 31-32, 48-49. See generally L. COHEN, *supra* note 363.

372. See *supra* text accompanying notes 296-303.

tive explanations of how it actually occurred, he has established only that the tortious condition contributed to the risk, not to the actual occurrence of the injury. He therefore should be able to recover from the defendant only if exposure to the risk is itself considered to be a legal injury, and his damages should be measured accordingly.³⁷³

In the alternative-causation cases, the problem is significantly different. The *Summers* problem does not arise simply because each defendant is equally likely (or unlikely) to have been the one that fired the shot that wounded the plaintiff. It arises because we are trying to use pure or "naked" statistical probabilities, not based on any causal generalization, to determine what actually happened. Thus, the problem exists in the DES cases even though one defendant may have controlled more than fifty percent of the market and is therefore most likely, from a statistical standpoint, to be the defendant who caused any particular injury.³⁷⁴ Naked statistical data generates very little confidence in deciding who actually caused the injury. For example, few people would be willing to say that Joe caused Sue's death simply because Joe fired seven bullets in her direction while Mike only fired three. No one would treat Joe as the cause, no matter how many more bullets he fired than Mike, if the bullet that killed her were identified by its markings as one that came from Mike's gun.

As evidence fitting a certain causal generalization builds up, it generates an increasingly strong belief in the existence of the asserted causal connection.³⁷⁵ The markings on the bullet are "particularistic" empirical data that fit into a causal generalization and therefore have explanatory force. On the other hand, naked statistical reports of who fired the most shots in Sue's direction do not fit into any causal generalization and thus do not generate any such belief.³⁷⁶ They are useful for placing a bet on

373. See *supra* text accompanying notes 312-29 & 334-41.

374. Eli Lilly & Co. may well have been in this position. Levine, *Gilding the Lilly*, TRIAL, Dec. 1984, at 19-20. The *Sindell* court carefully noted that an inference of causation by any one DES manufacturer based on probability would have failed "if we measure[d] the chance that any one of the defendants supplied the injury-causing drug by the number of possible tortfeasors." *Sindell*, 26 Cal. 3d at 603, 607 P.2d at 931, 163 Cal. Rptr. at 139 (emphasis added); see also *id.* at 611-12, 607 P.2d at 936-37, 163 Cal. Rptr. at 144-45. Implicitly, a probabilistic "more likely than not" inference would have been possible against Eli Lilly & Co. if the chance were measured by market shares rather than by the number of defendants, but, contrary to David Kaye's suggestion, this was not what the court was after. See Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, 1982 AM. B. FOUND. RESEARCH J. 487, 508, 514. Rather, the court was unwilling to hold a defendant liable for the entire damage based simply on naked statistical inference. Kaye's article misses a critical element when it treats proof as a simple matter of probabilities and ignores the need to induce the belief that a causal relationship actually existed.

375. L.J. COHEN, *supra* note 363, *passim*.

376. It is this distinction that explains the difference between Judith Jarvis Thomson's "external" evidence (naked statistical report) and her "internal" evidence (particularistic evidence that fits into a causal generalization). Thomson, *Remarks on Causation and Liability*, 13 PHIL. &

who caused Sue's death, but they do not help us to determine whether Joe or Mike actually was the cause.³⁷⁷ At most, they tell us the relative amounts of ex ante risk that Joe and Mike imposed on Sue, when they are combined with the causal generalization that links firing a gun in someone's direction with killing that person. Thus, they can be used to allocate liability that is based on exposing someone to the risk that was realized. Nevertheless, they are of little use in establishing who actually caused Sue's death.

The courts generally have perceived the critical distinction between naked statistical evidence and causally relevant particularistic evidence. For example, in *Smith v. Rapid Transit, Inc.*,³⁷⁸ the court refused to allow the jury to infer that one of the defendant's buses caused a particular accident when the evidence merely showed that the defendant's buses were the most frequent users of the route on which the accident occurred. Unless the statistical evidence is so compelling that it precludes any possible alternative explanation, it will be insufficient by itself to induce the belief that the asserted causal connection actually existed.

In sum, so long as tort liability continues to be based on individual responsibility, liability will be imposed on a defendant only if it is believed that the tortious aspect of his conduct actually contributed to the specified legal injury. This belief is the essence of a causal explanation, as distinguished from mere probabilistic statements of increased risk. The belief will arise in a particular case only if there is sufficient evidence that not only supports the suggested causal explanation involving the defendant's tortious conduct, but also makes it the most plausible suggested explanation. Ordinarily, only particularistic evidence fitting the relevant causal generalizations can accomplish this task.

CONCLUSION

Although causation intuitively seems to be a rather simple concept, it is in fact a complex and subtle one that long has resisted efforts to articulate a precise definition. The repeated failure of such efforts during

PUB. AFF. 101, 127-33 (1984); see also Nesson, *supra* note 363, at 1377-79, 1383-85 (acceptable-stories theory).

377. Many writers have confused the betting odds that a person is willing to accept on the existence of a certain fact with the belief that the person actually has in the existence of that fact. E.g., Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L.J. 101, 105; Tribe, *supra* note 363, at 1346-48; Wagner, Book Review, 1979 DUKE L.J. 1071, 1072-73 & n.6. As the text indicates, a willingness to accept betting odds does not necessarily imply *any* belief in the actual existence of the fact in question. Jonathan Cohen has written extensively and persuasively on the inductive nature of causal belief and the inapplicability of mathematical probability theorems to such belief. L. COHEN, *supra* note 363, summarized in Schum, *A Review of a Case Against Blaise Pascal and His Heirs*, 77 MICH. L. REV. 446 (1979); see T. BEAUCHAMP & A. ROSENBERG, *supra* note 21, at 312-14; Nesson, *supra* note 363, at 1385-90.

378. 317 Mass. 469, 58 N.E.2d 754 (1945).

this century has led to frequent confusion of the causal and noncausal issues in tort scholarship. More recently, it has led to a general denigration of the significance and even of the meaning of the concept of causation in tort law.

In this Article, I have argued that the concept of causation has a definite substantive content that was first elaborated by the philosophers David Hume and John Stuart Mill. I have tried to demonstrate the analytical clarity and power that results from a proper understanding of the concept of causation and its connection, through the tortious-aspect causation requirement, to the concept of legal responsibility. A voluminous literature that confuses the two concepts can be sorted out and integrated. Problematic cases can be solved. Restrictions on naked statistical evidence can be explained and justified. The policy issues posed by recent cases on the frontiers of tort liability can be correctly identified.

I have concentrated on these descriptive issues and problems. However, this Article also has important normative implications. The denigration of the concept of causation during this century has seriously eroded the traditional view of tort liability as a system of corrective justice, which protects individual autonomy by redressing injuries if and only if they were caused by the tortious aspect of the defendant's conduct. Tort scholarship instead has been dominated by books and articles that view tort liability as a (poor) system for providing general compensation for losses³⁷⁹ or as a (poor or great) system for maximizing social wealth.³⁸⁰

But the concept of corrective justice embodied in the tortious-aspect causation requirement clearly continues to control the decision of actual tort cases. Despite more than a half-century of academic assaults on the causation requirement, the courts almost invariably reach results that are consistent with it. The requirement and the underlying concept of corrective justice exert a very strong normative pull, even when judges do not articulate that pull but rather couch their decisions in the currently fashionable language of loss-spreading and wealth-maximization.

As I have noted elsewhere,³⁸¹ the concept of corrective justice has fallen on hard times in recent years because of difficulties encountered in defining its two principal elements, causation and rights, rather than from any normative or descriptive weakness in the concept itself, which continues to be both more normatively appealing and more descriptively

379. E.g., R. KEETON & J. O'CONNELL, BASIC PROTECTION FOR THE TRAFFIC VICTIM 1-3, 11-75, 249-50, 261-65 (1965); England, *supra* note 11, at 28-29, 62, 67-69. See generally *Symposium: Alternative Compensation Schemes and Tort Theory*, 73 CALIF. L. REV. 548 (1985).

380. E.g., G. CALABRESI, *supra* note 100, at 239-87 (poor system); Landes & Posner, *The Positive Economic Theory of Tort Law*, 15 GA. L. REV. 851 (1981) (great system).

381. Wright, *supra* note 9, at 435-36, 455-56.

accurate than any alternative theory of liability. This Article has defined and elaborated the causal element in the concept of corrective justice. In another article, now in progress, I address the rights element by elaborating a principled theory of tortious conduct. I hope, through both articles, to make explicit the principles that have provided the foundation for tort liability over the centuries, but which require more detailed elaboration in these skeptical times.