Causation: Linguistic, Philosophical, Legal and Economic

Richard W. Wright
IIT Chicago-Kent College of Law

Ingeborg Puppe
Rheinische Friedrich-Wilhelms University of Bonn

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

Part of the Law Commons

Recommended Citation
CAUSATION: LINGUISTIC, PHILOSOPHICAL, LEGAL AND ECONOMIC †
RICHARD W. WRIGHT* & INGEBORG PUPPE**

I. INTRODUCTION

Causation of a legally recognized injury by the wrongful aspect of the defendant’s conduct is a fundamental requirement, as a matter of interactive (“corrective”) justice and actual practice, for the defendant’s legal responsibility for such injury to an individual (through tort law for the private wrong) or to all members of society (through criminal law for the public wrong). To achieve rough or second-best justice in tort law in situations in which causation may exist but generally is impossible to prove, some courts have created presumptions of causation or imposed proportional liability based on a statistical probability of causation, but liability is never imposed if causation is clearly lacking, unless denial of liability would distort a second-best just liability scheme.2

†Copyright 2016 Ingeborg Puppe and Richard W. Wright, either of whom may grant permission to copy. All rights reserved. Permission is hereby granted to copy for noncommercial purposes as long as proper citation is made to this publication. This paper was written at the same time as two other papers on the same topics, portions of which contain substantially the same text as portions of this paper. See Ingeborg Puppe & Richard W. Wright, Causation in the Law: Philosophy, Doctrine and Practice, in THE COMMON CORE OF EUROPEAN PRIVATE LAW: CAUSATION (Marta Infantino & Eleni Zervogianni eds., 2016); Richard W. Wright, The New Old Efficiency Theories of Causation and Liability, 7 J. TORT L. 65 (2015) [hereinafter Wright, New Old Efficiency Theories]. We are very grateful to Thomas Grosse-Wilde for his extensive assistance and advice, without which none of our joint work would be possible.

*University Distinguished Professor and Professor of Law, Chicago-Kent College of Law, Illinois Institute of Technology.

**Emeritus Professor of Law, Rheinische Friedrich-Wilhelms University of Bonn.


2. See Wright, Possible Wrongs, supra note 1, at 1325–30 (discussing Hymowitz v. Eli Lilly & Co., 539 N.E.2d 1069 (N.Y. 1989)), in which the court refused to let defendants who marketed DES for birth-related purposes avoid liability by proving lack of causation for a specific plaintiff since allowing such would disrupt the proportional market-share liability adopted as a second-best just liability scheme).
In this paper, we briefly survey the linguistic, philosophical, legal, and economic usages and analyses of the concept of causation. In part II, we discuss the ambiguous usages of causal language in ordinary speech and legal discourse and argue that, to promote clear identification and proper analysis of the basic natural/factual/actual causation issue and the distinct normative legal responsibility issue, such ambiguity should be eliminated in legal discourse by using causal language to refer solely to causation in its basic sense. In part III, we discuss the philosophical foundations of the modern analysis of causation in its basic sense. While lawyers frequently scoff at philosophical analyses of causation, their failure (shared by many philosophers) to clearly distinguish the descriptive issue of causation in its basic sense from the normative issue of legal responsibility and to appreciate John Stuart Mill’s path-breaking “covering law” analysis of causation in its basic sense has until now mired law in considerable confusion. In part IV.A, we criticize the strong necessity (sine qua non or “but for”) criterion, especially in its usually assumed counterfactual form, which is often claimed to be the exclusive criterion for identifying causes in specific situations. In part IV.B, we discuss the more comprehensive weak necessity/strong sufficiency criterion, based on Mill’s covering law analysis, of which the strong necessity criterion is merely a corollary that works only when there is no causal overdetermination. In part V, we discuss the procedural standards of persuasion, which should and usually do require formation of a minimal belief rather than a mere statistical probability, and the courts’ adoption of second-best liability rules when it is inherently impossible to prove or disprove causation. In part VI, we criticize the legal economists’ attempts to explain the causation requirement for legal responsibility, despite causation’s being irrelevant under the efficiency theories of liability. Part VII contains a brief conclusion.

3. See, e.g., RESTATEMENT (SECOND) OF THE LAW OF TORTS § 431 cmt. a (AM. LAW INST. 1965) (distinguishing the legal and popular sense of causation from “cause in the so-called ‘philosophic sense’”); FREDERICK POLLOCK, THE LAW OF TORTS: A TREATISE ON THE PRINCIPLES OF OBLIGATIONS ARISING FROM CIVIL WRONGS IN THE COMMON LAW (11th ed. 1920) ("the lawyer cannot afford to adventure himself with philosophers in the logical and metaphysical controversies that beset the idea of cause"); Spier & Haazen, supra note 1, at 130–31; Jane Stapleton, CAUSATION IN THE LAW, in THE OXFORD HANDBOOK OF CAUSATION 744, 749 (Helen Beebee et al. eds., 2009) [hereinafter OXFORD HANDBOOK] ("Traditionally, most lawyers disdain philosophical inquiries into causation. In my opinion, this indifference is warranted.").
II. CAUSAL TERMINOLOGY: DISTINGUISHING CAUSATION PER SE FROM LEGAL RESPONSIBILITY

In ordinary speech as well as legal discourse, attributions of causation usually are based not only on causation in its basic sense, which refers to the content and operation of the laws of nature (“causal laws”), but also to identify one or more of the contributing conditions as being more significant than the others in the particular context. In law, the usual relevant context is the assessment of legal responsibility for some specific state of affairs. Thus, in legal discourse, causal language is often employed to refer not only to causation in its basic sense, but also—employing phrases such as “lien de causalité,” “Kausalzusammenhang,” “legal cause,” “proximate cause,” “material contribution,” “substantial factor,” or merely “cause”—to the normatively based principles for limiting the defendant’s legal responsibility for the consequences of his wrongful conduct.

The principal limitations on legal responsibility for wrongfully caused consequences also invoke causal laws. They do so, however, not with reference to the causal status of the defendant’s wrongful conduct, which has already been found or assumed to be true, but rather with reference to whether the legal injury (and related damages) (i) almost certainly would have occurred anyway as a result of some other, non-responsible condition(s) (the “no worse off” limitation, which is often confused with the conditio sine qua non (“but for”) test for causation per se), (ii) occurred only because of the intervention of some other, highly unexpected cause (the “superseding cause” limitation, including “force majeure”), or (iii) did not result from the continued operation or playing out along the causal chain of a risk that made the conduct wrongful (variously described as the “adequacy,” “scope of the risk,” “risk playout” or “continuity” limitation).


6. For discussion of the no worse off, superseding cause and risk playout limitations, see Wright, Legal Responsibility, supra note 1, at 1434–51; cf. Spier & Haazen, supra note 1, at 141–46 (no worse off situations). For discussion of the continuity limitation, see Ingeborg Puppe, Negligence and Responsibility in German Road Traffic Law, 11 Eur. J. Crime, Crim. L. & Crim. Justice 151, 161–62 (2003) [hereinafter Puppe, Road Traffic]; cf. Fairgrieve & G’sell-Macrez, supra note 1, at 119 (discussing Dejean de la Bâtie’s “continued spread of evil” analysis, which requires “that when many events occur between the initial act and the damage, each of these facts must contain an element of unlawfulness that
The use of causal terminology to refer not only to causation in its basic sense but also to the normative issue of attributing legal responsibility has generated considerable disagreement and confusion over the nature of and differences between these two quite different issues and, as a result, frequent failures to distinguish them and to apply the proper empirical or normative analysis, respectively, to each of them.7 As the American Law Institute (ALI) has recently emphasized, legal reasoning by courts and lawyers would be much clearer and properly focused if causal terminology were applied only to causation in its basic sense. After having promoted the confusion of causation and responsibility for 75 years by lumping the empirical and normative issues together under the phrases “legal causation” and “substantial factor” in the first and second Restatements of the Law of Torts,8 the ALI in the recently published Restatement Third uses “factual causation” to refer to causation in its basic sense and “scope of liability” to refer to the separate issue of appropriate legal responsibility.9 In the remainder of this paper, we focus solely on the issue of causation in its basic sense.

III. PHILOSOPHICAL FOUNDATIONS

A. David Hume’s Regularity Theory

Writing in the first half of the 18th century, David Hume argued, contrary to the then popular belief, that singular causal judgments are not based on direct perception of causation or causal qualities or forces inherent in objects or events. No such directly perceptible qualities or forces have ever explains the unlawfulness of the subsequent fact”). For discussion of the adequacy and scope of the risk limitations, see Cees Van Dam, European Tort Law 312–16 (2d ed. 2013); Van Gerven et al., supra note 5, at 397–407; 2 Christian Von Bar, The Common European Law of Torts 439 (2000); Fairgrieve & G’sell-Macrez, supra note 1, at 118–19, 120–21; Spier & Haazen, supra note 1, at 132–33.


9. Restatement (Third) of the Law of Torts: Liability for Physical and Emotional Harm §§ 26, 27 & 29 (Am. Law Inst. 2010). However, confusion persists in the Restatement Third’s elaboration of the substantive criteria and procedural standards of proof for resolving the factual causation issue. See Wright, Possible Wrongs, supra note 1, at 1311–12, 1339–40; infra note 113.
been identified. Instead, Hume insisted, all we observe are uniform regularities of succession, whereby the occurrence of one object is invariably followed by the occurrence of a temporally and physically contiguous distinct object, from which we psychologically infer or induce a necessary connection between the first object and the second.\textsuperscript{10}

Hume provided two definitions of a cause and the related idea of a necessary connection, which he seemed to view as being equivalent, but which are quite different:

\[ \text{We may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second. Or in other words where, if the first object had not been, the second never had existed.}\textsuperscript{11} \]

The first definition describes the occurrence of the first object (the cause) as being sufficient for (necessitating) the occurrence of the second object (the effect). The second defines the occurrence of the first object as being necessary for the occurrence of the second. Moreover, it assumes that the first object (or one of its properties) is strictly necessary for the occurrence of the second, in the sense that the second never occurs in the absence of the first (or some property of the first shared by all other causes of the second). Hume’s belief that these definitions are equivalent reflects his apparent belief that the causal relation is a two-way (if and only if) necessity relation: neither object (as a whole or with respect to some property) can occur in the absence of the other.\textsuperscript{12} His discussion focused on the first definition, which has come to be known as the “regularity” account and is subject to numerous counter-examples.\textsuperscript{13} Philosophers and lawyers, therefore, have generally focused on the second (necessary condition) definition, although in a less strict sense.

\begin{itemize}
\item \textit{David Hume, A Treatise of Human Nature} bk. I, pt. III, § XIV (1739–1740) [hereinafter \textit{Hume, Treatise}]; \textit{David Hume, An Enquiry Concerning Human Understanding} § VII, pts. I–II (1748) [hereinafter \textit{Hume, Enquiry}]. Sandy Steel claims, like Mackie but contrary to Hume and almost every other philosopher, that we sometimes can directly perceive causation without even implicitly relying on assumed uniform regularities or causal laws. \textit{Steel, supra} note 1, at 67, 67 n.107; see infra text accompanying note 34.
\item \textsuperscript{10} \textit{Hume, Enquiry}, supra note 10, § VII, pt. II.
\item \textsuperscript{11} \textit{Hume, Treatise}, supra note 10, bk. I, pt. III, § XV (rules 4–6).
\item \textsuperscript{12} \textit{See, e.g., Lewis, supra} note 4, at 563–67. Incredibly, von Bar states that whether a distinction between “the causal and the merely temporal succession of events . . . is fundamentally justifiable is irrelevant.” \textit{Von Bar, supra} note 6, at 435–36.
\end{itemize}
In his doctoral thesis, *On the Fourfold Root of the Principle of Sufficient Reason*, originally published in 1813 and substantially revised in 1847, Arthur Schopenhauer identified the Law of Causality as one of the four different applications of the Principle of Sufficient Reason, which states that nothing is without a reason for its being. The four roots are distinguished from one another based on the nature of the relation at issue.  

The Law of Causality, which applies to physical relations among material objects, is the principle of sufficient reason of becoming, which states an empirical physical necessity for the occurrence of specific changes in some states of affairs given prior changes in other states of affairs. Although he did not refer to Hume’s analysis in the original 1813 publication, he agreed with Hume’s empirical inductive approach to determining causal relations, but as an a priori truth about material reality rather than a mere psychological belief. In a section added to the 1847 edition, he curtly dismissed Hume’s apparent skepticism about any real laws of causation. Moreover, contrary to Hume’s apparent supposition, Schopenhauer observed that the cause of some effect is not a single prior object or changed condition but rather a set of conditions, each of which is necessary for the set to be sufficient for the change in a different condition that constitutes the effect. While we often designate as the cause the last change to occur that completes the sufficient set, scientifically and philosophically, the cause is the entire set.  

**C. John Stuart Mill’s Laws of Nature (Covering Law) Theory**

In *A System of Logic*, first published in 1843, John Stuart Mill corrected and advanced Hume’s analysis on numerous points, most of which are ignored by current philosophers. Like Schopenhauer, Mill replaced Hume’s skeptical treatment of causation as a mere psychological belief with an account of causation as the instantiation in specific situations of the laws of nature. He explained that to be a law of nature—a *causal* law—the lawful (always true) relation must be empirically induced from observation and

15. *Id.* § 20.  
16. *Id.* § 12.  
17. *Id.* § 20.  
18. For ease of reference we will use this term to refer to Mill’s account, although many covering law and related deductive nomological theories fail to capture all of the details of Mill’s theory or its subsequent elaborations and extensions.
reflection on natural events and states of affairs, rather than being merely analytical—that is logical, definitional, mathematical, or mereological (the relation of a whole to its constituent parts). Moreover, to be a law of nature, a regularity has to be not only invariable in experience but also unconditional—that is, not subject to any exceptions under possible actual conditions. It also must be a member of the set of the simplest, fewest, mutually consistent such regularities from which all causal generalizations, natural events, and states of affairs can be explained.19 Although philosophers generally fail to recognize or acknowledge this, these conditions render Mill’s laws of nature account immune to the counter-examples that plague Hume’s regularity theory.20

Like Schopenhauer, Mill also differed from Hume by noting that the antecedent portion of a causal law rarely, if ever, includes only a single abstract object, but instead consists of a minimally sufficient set of abstract conditions which, when fully instantiated in a specific situation, is sufficient for the instantiation of the consequent of the causal law. Like Schopenhauer, he insisted that, although we usually refer to only one or some of the instantiated conditions in the antecedent as the cause(s), with the chosen condition(s) varying depending on our purpose in the specific context, and treat the others as mere “conditions,” philosophically and scientifically there is no basis for such discrimination.21 He described the proper scientific methods of observation and experimentation, including especially the Method of Difference (experimentally changing only one condition while holding all others constant), for attempting to determine by induction and ratiocination the minimally sufficient conditions in the antecedent of a causal law.22

Despite stating that “from nothing, a mere negation, nothing can proceed,” Mill noted that the antecedent of a causal law will almost always include not only positive conditions but also negative conditions, not as “producing” causes but as “the mere absence of a preventing cause,” the absence of which allows a positive causal process that otherwise was being or would have been prevented to proceed.23

22. Id. bk. III, chs. VII–VIII.
23. Id. bk. III, ch. V, § 3, vii n.*; contra, e.g., Moore, supra note 7, at 53–55, 399–400, 436–37, 444–45, 460–61; Roderick Bagshaw, Causing the Behaviour of Others and Other Causal Mixtures, in PERSPECTIVES ON CAUSATION, supra note 1, 361, 374–75; Douglas Ehring, Causal Relata, in OXFORD HANDBOOK, supra note 3, 387, 397–98. Omissions and other absences can be producing as well as
Mill also noted, contrary to Hume’s assumption, that more than one abstract set of minimally sufficient conditions may be specified for the occurrence of a certain effect—for example, death, which can occur in different ways—and that different causes may interact in a specific situation to reinforce or counteract each other. However, he assumed (i) that only one minimally sufficient set would be instantiated in any specific situation and (ii) that the joint operation of independent causal processes in a specific situation would always produce dissimilar effects than those that would be produced by each acting separately—either a summation of the effects of each acting separately or a completely different effect. Thus, although he rejected Hume’s strict necessity criterion for a causal condition, according to which a condition is not a cause of some effect unless the effect never occurs in the absence of that condition, he apparently assumed that a condition must be strongly necessary, in the sense that it was necessary for the occurrence of the effect in the specific situation. While this assumption may be true when considering whole events or states of affairs, at least in cases of duplicative rather than preemptive causation, it is false when considering the properties of events or states of affairs, which are the scientifically, philosophically, and legally relevant causal relata.

Mill explained that, although mathematical laws or theorems—for example that two plus two equals four or that the angles of a triangle always add up to 180 degrees—are not empirical laws of succession and thus are not in themselves causal laws, as general laws of number and space they often are included as analytical elements in causal laws. Similarly, in jurisprudence, we have to use institutional rules created by humans, such as voting and liability rules, to determine legal results. These institutional rules are not causal laws, or indeed laws at all in the scientific or philosophical

allowing causes and are essential elements in every causal process involving human actions and many not involving human actions. See Dieter Birnbacher & David Hommen, Omissions as Causes—Genuine, Quasi, or Not at All?, in CRITICAL ESSAYS ON “CAUSATION AND RESPONSIBILITY”, supra note 5, 133; Ingeborg Puppe, Der Erfolg und seine kausale Erklärung im Strafrecht, 92 ZStW 863, 895–99 (1980) [hereinafter Puppe, Der Erfolg] (English translation available at http://papers.ssrn.com/abstract_id=2743259); Puppe, Concept, supra note 5, at 82–83; Jonathan Schaffer, Causes Need Not Be Physically Connected to Their Effects: The Case for Negative Causation, in CONTEMPORARY DEBATES IN PHILOSOPHY OF SCIENCE 197 (Christopher Hitchcock ed., 2004); Wright, NELL Account, supra note 20, at 311–21; Richard W. Wright, Causation: Metaphysics or Intuition, in LEGAL, MORAL, AND METAPHYSICAL TRUTHS: THE PHILOSOPHY OF MICHAEL S. MOORE 171, 177–82 (Kimberly Kessler Ferzan & Stephen J. Morse eds., 2016). The cited papers by Wright provide the metaphysical account of negative conditions as causes and the criticism of opposing accounts that Stapleton claims that he fails to provide. See Stapleton, Extended But-For, supra note 7, at 701–02.


25. See infra text accompanying notes 56–60.

tical sense of being universally applicable, but they are applicable in a cer-
tain society and they are general regularities, which like causal laws govern
single cases and can be used to explain them. However, although institutional
rules govern institutional facts independent of their psychological
effects, the bare fact of their being satisfied in a particular case has no natu-
ral/actual effect, which occurs only if and when the institutional fact of
their being satisfied is recognized by the relevant parties and/or officials
and causes them to take certain actions, as with the concrete application of
mathematical rules in specific instances in commerce and life.27

As Hume emphasized with regard to assumed uniform regularities,28
Mill emphasized that our knowledge of the laws of nature, being inductive-
ly derived from actual experience, can never be assumed to be complete.
Even if it were complete, we ordinarily employ causal generalizations,
which are incomplete (and thus contingent) and encompass, usually at a
gross macro level, a multitude of successive or simultaneously operative
more specific generalizations and the underlying laws of nature.29 Never-
theless, he insisted, an assertion of causation always involves, implicitly,
an assertion of the complete instantiation of a network of underlying causal
laws, even when this assertion is based on a single observation.30

D. John Mackie’s INUS / Strong Necessity Analysis

Among philosophers, the best known Millian analysis of causation
was initially published in 1965 by John Mackie, who employed an acro-
nym, INUS (for “insufficient but non-redundant [necessary] part of an un-
necessary but sufficient condition”), to refer to the conditions that make up
the minimally sufficient set of abstract conditions that constitute a causal
law.31 If applied as a criterion for being an actual causal condition in a con-
crete singular instance, the INUS criterion would be one way of describing
the least stringent, weak sense of necessity, which merely requires that a
condition be necessary for the sufficiency of a set of actual conditions that
was sufficient for the occurrence of the effect, rather than being always

27. Thus, the use of institutional rules as part of the analysis of causation in law and other areas of
life does not expose any deficiency in the laws-of-nature based covering law account of causation.
Contra STEEL, supra note 1, at 32–33, 35; Stapleton, Extended But-For, supra note 7, at 699, 702 n.19,
703; cf. Puppe, Concept, supra note 5, at 84–85.
28. HUME, TREATISE, supra note 10, bk. I, pt. III; HUME, ENQUIRY, supra note 10, § IV, pt. II; §
V, pt. I; § VII, pt. II.
31. MACKIE, supra note 4, at 62. Chapter 3 (pages 59–87) of MACKIE is a revised version of
necessary \((\text{strict necessity})\) or necessary for the effect in the singular instance \((\text{strong necessity})\).\(^{32}\) However, Mackie and other philosophers writing around the same time followed Mill in employing this weak necessity criterion only to identify the minimally sufficient set of conditions in a causal law, while insisting on the strong necessity criterion in singular instances of causation.\(^{33}\) Furthermore, unlike Mill, Mackie claimed that the strong necessity criterion can be and often is applied in specific instances using singularist analogical/contrastive reasoning, without any need to refer even implicitly to causal laws.\(^{34}\) Both philosophers and lawyers generally fail to note these aspects of Mackie’s theory and credit him, rather than Herbert Hart and Tony Honoré, who wrote before him and of whose work he was aware,\(^ {35}\) with the initial development and application of the weak necessity criterion for singular instances of causation.

\section*{E. David Lewis’s Counterfactuals Theory}

In both philosophy and law, it is generally assumed that application of the strong necessity criterion requires a counterfactual (hypothetical) inquiry: if the actual condition at issue had not occurred, what would have happened?\(^ {36}\) However, logicians point out that this counterfactual inquiry cannot produce any determinate truth value. If the antecedent of a condition is false, this so called unreal conditional clause is true whether the consequent is true or not. \textit{Ex falso quodlibet}.\(^ {37}\)

In an attempt to provide a logically valid basis for the counterfactual interpretation of the strong necessity criterion, David Lewis introduced possible worlds, which he considered to be real, albeit slightly different


from our world. In such a possible world the statement that the actual condition that occurred in our real world did not occur can be true, so the statement that in this possible world the consequence also did not occur makes sense. Lewis treats an event as causal for a result if the possible world in which this event and its result do not occur is more similar to our real world than any other possible world. To determine what would happen in various possible worlds, we have to presume that they are governed by causal laws, and if they are similar to our real world, these causal laws have to be the same as in our real world. But in such a possible world you cannot omit any event or fact which has taken or would take place according to those causal laws, given all the existing and prior conditions, without violating those laws. To solve this problem, Lewis arbitrarily stipulated that a targeted violation of the causal laws (i.e., a miracle) is less of a departure from our real world than a wholesale change in the past and future causal history. He relies on such targeted miracles to get the desired factual situation in which the condition at issue is not present in the possible world and then applies the causal laws.38

As the above description of Lewis’s possible worlds analysis indicates, the strong necessity criterion depends for its application on a Millian covering law analysis, although this is rarely acknowledged by proponents of the strong necessity criterion. The only rational way to answer the question whether a certain condition would have happened if some other condition had not happened is to refer to universal laws under which one can subsume the fictional case that the latter condition did not occur.39 Thus, almost all proponents of the strong necessity criterion are, unlike Mackie, causal generalists, in the sense that they explicitly or implicitly base the necessity of a condition for the occurrence of the effect on causal laws. In overdetermined causation situations, in which there were two or more minimal sets of conditions that were actually sufficient (duplicative causation) or hypothetically sufficient (preemptive causation) for the occurrence of the effect, the proponents of the strong necessity criterion often shift, without realizing it, into covering law analysis rather than strong necessity analysis. For example, Mackie states that we usually have no difficulty identifying


the actual cause in preemptive causation situations, but his explanation of why this is so relies on covering law analysis rather than strong necessity: “Where we have no hesitation in making causal statements we can tell some more detailed causal story . . . . But the rival story about the alternative or reserve cause [preempted condition] cannot be completed.” 40

However, for Lewis and his followers, who have dominated the philosophical discussion of causation for the last forty years, 41 the essence of singular instances of causation is not the subsumption of a specific situation under some set of causal laws or generalizations, but rather the counterfactual sentence about the most similar possible world with localized targeted miracles, assuming the same causal laws as in our real world. Setting aside the resort to miracles, asking what hypothetically would have occurred if the condition at issue had not taken place leaves the way open for indeterminate speculation. For example, if the driver had not been going ten miles over the speed limit, how fast or slow would he have been going, or would he instead have been playing golf? The practical solution to this problem in the law, as in Lewis’s possible worlds theory using targeted miracles, is to reverse or think away only the condition at issue and then to run the causal laws, to the extent possible, to try to figure out what would have happened, while ignoring the fact that reversing the condition at issue would require changing many prior conditions in the possible world given those causal laws, which changes would result in quite different effects when running the causal laws. 42 Yet, even if we do this, we are often left to indeterminate speculation regarding what would have occurred once the condition at issue is reversed or thought away. 43

As Leon Green long ago noted, the counterfactual interpretation of the strong necessity criterion “take[s] the eye off the ball” by asking what would have occurred if things had been different, rather than how what happened did occur given things as they actually were. 44 The strong neces-

40. Mackie, supra note 4, at 45; see id. at 33–34 (relying on covering law analysis to resolve the epiphenomena problem); Honoré, supra note 35, at 376–79; Jane Stapleton, Choosing What We Mean by “Causation” in the Law, 73 Mo. L. Rev. 433, 452–53, 453 n.45 (2008) [hereinafter Stapleton, Choosing]; Stapleton, Extended But-For, supra note 7, at 709 (“the prevention mechanism . . . would have been complete”). For elaboration of the distinction between duplicative and preemptive causation, see infra text at note 49.

41. E.g., CAUSATION AND COUNTERFACTUALS (John Collins et al. eds., 2004).

42. E.g., Steel, supra note 1, at 17; Friedrich Toepel, Causal Overdetermination, in CRITICAL ESSAYS ON “CAUSATION AND RESPONSIBILITY,” supra note 5, 111, 125–26.

43. Moore, supra note 7, at 85; Puppe, Der Erfolg, supra note 23, at 868–70; Puppe, Road Traffic, supra note 6, at 151–53; Schaffer, Contrastive Causation in the Law, supra note 36, at 270–71, 283–85; Wright, Pruning, supra note 32, at 1040–41.

sity criterion, when properly applied through a real world factual analysis rather than a possible worlds counterfactual analysis, is a corollary of the covering law analysis of causation that is valid only when there is no causal overdetermination\(^45\) and has two steps, the second of which is causally irrelevant but is satisfied when there is no causal overdetermination: (1) was the condition at issue part of the complete instantiation in the specific situation of the antecedent of one or more relevant causal generalizations and their underlying laws that have as their ultimate consequence the effect at issue (the covering law analysis); and (2) were the other existing conditions insufficient without the condition at issue for such complete instantiation (the strong necessity analysis)?\(^46\)

IV. CAUSATION IN THE LAW

A. Strong Necessity (Sine Qua Non / But For)

In both law and philosophy, the usually assumed criterion for identifying a condition as a cause in a specific situation is the strong necessity (\textit{sine qua non}) criterion, which states that a condition was a cause of an effect if the effect would not have occurred in the absence of (but for) the existence of the condition in the specific situation.\(^47\) As was discussed in part III.E immediately above,\(^48\) it also is generally (but erroneously) assumed that proper application of the strong necessity criterion—or any other necessity criterion—requires a counterfactual, hypothetical inquiry: if the actual condition at issue had not occurred, what would have happened?

Even if the strong necessity criterion is properly interpreted as a factual inquiry into what actually happened, based on instantiation of the laws of nature by the actual conditions in the specific situation, rather than a coun-

\(^{45}\) We use “overdetermination” literally to include situations involving preemptive causation as well as duplicative causation. See infra text at note 49; cf. Lewis, supra note 4, at 567, 567 n.12 (same usage).


\(^{47}\) E.g., DIGEST OF EUROPEAN TORT LAW: ESSENTIAL CASES ON NATURAL CAUSATION 99–100 (Benedict Winiger et al. eds., 2007) [hereinafter DIGEST]; DAN B. DOBBS, THE LAW OF TORTS 409 (2000); VAN DAM, supra note 6, at 310; MACKIE, supra note 4, at 38–48, 76–77, 126–27; RESTATMENT (THIRD) OF THE LAW OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 26 (AM. LAW INST. 2010); VON BAR, supra note 6, at 435 n.1; Fairgrieve & G’sell-Maurez, supra note 1, at 117; 120, 129; Lewis, supra note 4, at 557–58; Olivier Moréteau, French Tort Law in the Light of European Harmonization, 6 J. CIV. L. STUD. 768, 769 (2013) (noting the proposed use of the strong necessity criterion in the Terré draft of proposed revisions in the French civil code, which, like many codes, currently contains no criterion for establishing actual causation); Spier & Haazen, supra note 1, at 127.

\(^{48}\) See supra text accompanying notes 36–46.
terfactual inquiry about what might have happened if things had been different, it fails as a comprehensive criterion for identifying causes. It cannot properly resolve situations involving duplicative or preemptive causal overdetermination. In duplicative causation situations, there are two or more distinct (although usually overlapping) fully instantiated minimally sufficient sets of conditions, for example, two fires each sufficient without the other but in combination with other conditions—such as a house to burn, oxygen, a specific wind direction, sufficient fuel on the path between the origin of the fire and the house, and lack of adequate countervailing rain or fire department efforts—to destroy a house, which merge and destroy the house. In preemptive causation situations, the preemptive cause is part of a fully instantiated minimally sufficient set while the preempted condition was a member of an incompletely instantiated minimally sufficient set. For example, if the first fire destroys the house before the second fire arrives, one of the necessary conditions for a minimally sufficient set that includes the second fire (the presence of a house to burn down when the second fire arrives) did not exist. In either situation, contrary to the laws of nature, common sense, and the decisions of the courts, neither fire would be treated as a cause under the strong necessity criterion, since neither was strongly necessary given the existence of the other.

Contrary to what proponents of the strong necessity criterion as the exclusive criterion often assume, instances of overdetermined causation are not rare. Some of the proponents of the strong necessity criterion treat the attributions of legal responsibility in overdetermined causation cases as policy based departures from the usual requirement of a causal connection between the wrongful aspect of the defendant’s conduct and the plaintiff’s injury. Others have tried to modify

49. HART & HONORÉ, supra note 4, at 122–28, 206–07, 235–53; Puppe, Der Erfolg, supra note 23, at 863, 868–70; Puppe, Concept, supra note 5, at 69–70, 78–80; Wright, Causation, supra note 46, at 1775–76.


51. E.g., MODEL PENAL CODE § 2.03 cmt. 2 at 259–60 (AM. LAW INST. 1985); VON BAR, supra note 6, at 440–41; Franz Bydlinski, Causation as a Legal Phenomenon, in CAUSATION IN LAW 7, 20–21 (Luboš Tichý ed., 2007).

52. See supra sources cited in note 50.

CAUSATION

the strong necessity criterion to handle these cases. However, as we explain in the remainder of this part, none of the modifications are successful.

The most frequent modification is to add details to the description of the injury, including not only the time and/or location at which it occurred but also tautological references to the causal process by which it occurred. For example, in the two-fires situation, the injury is described as the destruction of the house (i) at this particular time, (ii) with this particular debris pattern, and/or (iii) by two fires. By describing the injury in this way, all of the duplicative conditions become strongly necessary for its occurrence; alternatively, the preemptive cause but not the preempted condition becomes strongly necessary.\(^\text{54}\) Such techniques assume the causal conclusion, and then use that assumption to specify the desired level of detail of the injury or its manner of occurrence. They could be used to prove that anything contributed to the injury, no matter how causally irrelevant, merely by adding it to the description of the injury (e.g., that the injured person was wearing a hat), while ignoring the fact that the added details generally are irrelevant to the description of the required legal injury at issue.\(^\text{55}\)

The basic problem with these techniques is that they treat events or states of affairs as a whole, which can be described incompletely in almost infinitely variable ways, as the entities that are involved in a causal relation (the causal relata).\(^\text{56}\) If someone is asked to describe an event or state of affairs in its full concrete detail, she would have to describe the present criterion—albeit not sufficiently distinguished from the aggregate strong necessity criterion—to find that each individual offender’s possession of images of a child’s sexual abuse was a cause of the emotional and consequent economic harms suffered by the child due to her knowledge of widespread possession of those images).

\(^{\text{54.}}\) E.g., MODEL PENAL CODE § 2.03 cmt. 2 at 259; MACKIE, supra note 4, at 45–46; STEEL, supra note 1, at 18 (including only the time of occurrence); L.A. Paul, Counterfactual Theories, in OXFORD HANDBOOK, supra note 3, 158, 178–79; Spier & Haazen, supra note 1, at 128. For especially egregious examples of this approach, see Stapleton, Choosing, supra note 40, at 442 n.19 (death “by two bullets”), 452 (death “by electrocution at that instant” or “by explosion at 1 a.m.”); Stapleton, Extended But-For, supra note 7, at 700 (death “by poison at noon on Friday 13 June 2014 under Dan’s palm tree”), 704 (same), 710 (death “by a poison”), 711 (time, location, and “debris pattern”), 723–24 (time, location, and death “by thirst”). Stapleton claims that “the law knows” in hindsight what happened and how it happened, without explaining how “the law” knows this without relying on the covering law account of causation. Id. at 703–05, 710–11. Steel claims without elaboration that the NESS weak necessity account requires specification of the time of occurrence in preemptive causation situations. STEEL, supra note 1, at 18 n.16. It does not. See Wright, NESS Account, supra note 20, at 292, 297–303; Wright, Once More, supra note 5, at 1112–15.

\(^{\text{55.}}\) Puppe, Der Erfolg, supra note 23, at 870–74; Puppe, Concept, supra note 5, at 73–74, 79–80, 93–94; Schaffer, Contrastive Causation, supra note 7, at 302–05; Wright, Causation, supra note 46, at 1777–80.

\(^{\text{56.}}\) As Toepel now acknowledges. Toepel, supra note 42, at 117–22, 118 n.25.
state of the whole universe.57 Although many philosophers assume otherwise,58 causation occurs between specific properties of events and states of affairs, rather than between those events or states of affairs as a whole.59 In law, the required causal relation is between the wrongful aspect of the defendant’s conduct and the properly described legal injury, which usually does not include its specific timing or location. Instead those details serve merely to identify the specific event or state of affairs for which causation of the legally relevant properties (the required legal injury) is at issue.60

Friedrich Toepel, one of the principal defenders of the strong necessity criterion as the exclusive criterion, is perhaps the only one who recognizes that one must justify specifying some details of the relevant event or state of affairs (such as its timing or location) as part of the relevant legal injury while not also specifying others. Thus, he does not apply this approach in duplicative causation cases; instead he stands almost alone in heroically insisting that none of the duplicative conditions individually was a cause.61

He suggests that all supposed duplicative causation cases might turn out upon sufficiently detailed examination to be cases in which one condition

57. RUDOLF CARNAP, MEANING AND NECESSITY: A STUDY IN SEMANTICS AND MODAL LOGIC 29 (2d ed. 1956) (“If we require of a fact this maximum degree of completeness . . . , then there is only one fact, the totality of the actual world, past, present, and future.”).

58. See, e.g., Geert Kell, Making Causal Counterfactuals More Singular, and More Appropriate for Use in the Law, in CRITICAL ESSAYS ON “CAUSATION AND RESPONSIBILITY,” supra note 5, 157, 162; Lewis, supra note 4, at 558; Stathis Psillos, Regularity Theories, in OXFORD HANDBOOK, supra note 3, 131, 144–50; cf. PAUL & HALL, supra note 38, at 7, 7 nn. 1–3, who state that they do not agree that events are the proper causal relata, but nevertheless assume that they are in their subsequent analysis.

59. E.g., TOM L. BEAUCHAMP & ALEXANDER ROSENBERG, HUME AND THE PROBLEM OF CAUSATION 251–52, 255–56, 269–75, 281–82 (1981); MACKIE, supra note 4, at 256–58, 260–67; Ehring, supra note 23, at 406–07; Richard Fumerton, Moore, Causation, Counterfactuals, and Responsibility, 40 SAN DIEGO L. REV. 1273, 1278 (2003); Thomas Grosse-Wilde, Die Relata eines juristischen Kausalbegriffs und der juristische Syllogismus, in JUNGE RECHTSPHILOSOPHIE 43 (Carsten Bäcker & Sasha Ziemann eds., 2012); L.A. Paul, Aspect Causation, 97 J. PHI. 235 (2000), reprinted in CAUSATION AND COUNTERFACTUALS, supra note 41, 205; Wright, Pruning, supra note 32, at 1033–34, 1033 n.171. After a rigorous review of the various positions, Moore now concedes, contrary to his prior arguments, that concrete properties of events and states of affairs rather than events or states of affairs as a whole are the causal relata. Moore, supra note 7, at 361–65, 368 n.61. However, for reasons internal to his theory and relying on a misdescription of legal practice, he continues to treat events as the relevant causal relata. Id. at 366–68, 395–96; see Wright, NESS Account, supra note 20, at 287 n.9. Schaffer, like Moore, is forced to treat events as the causal relata by his contrastive account of causation. See Schaffer, Contrastive Causation, supra note 7, at 316–17.

60. Grosse-Wilde, supra note 59, at 49–50; Honoré, supra note 35, at 378–79; Puppe, Der Erfolg, supra note 23, at 878–82, 888–93; Puppe, Concept, supra note 5, at 83–84, 93–94; Puppe, Road Traffic, supra note 6, at 154; Wright, Causation, supra note 46, at 1759–74; Wright, NESS Account, supra note 20, at 287 n.9. Steel states that in Germany, if the plaintiff proves that the injury was caused by the defendant’s conduct as a whole, the burden is on the defendant to prove as a defense that the injury was not caused by the wrongful aspect of the conduct. STEEL, supra note 1, at 41.

61. Toepel, supra note 42, at 117, 119–20; see also Edelman, supra note 53, at 20, 23, 25–26; David Lewis, Causation as Influence, in CAUSATION AND COUNTERFACTUALS, supra note 41, 75, 80.
2016]  

CAUSATION 477

caused the relevant legal injury immediately before the other(s) would have done so and thus actually involve preemption rather than duplication. 62 He attempts to rescue the strong necessity criterion in preemptive causation cases by specifying the time at which the injury occurred. 63 He justifies including this detail in the specification of the legal injury by arguing that it is an essential aspect in cases of killing or other injuries to person or property, for which shortening life or usefulness allegedly is essential. 64

Both parts of this argument fail. With respect to duplicative causation, consider multiple independently sufficient forces, fires, doses of poison, etc. that combine prior to coming into contact with the adversely affected person or property. They clearly reinforce rather than preempt each other no matter how detailed the description may be. With respect to preemptive causation, conditions that cause death or other injuries may well do so at the same time or even later than the preempted condition would have done so, for example in Karl Engisch’s executioner hypothetical and the much discussed situation in which a traveler in the desert dies of thirst rather than poisoning when the poison put in his water canteen by one enemy is emptied out by another enemy who is unaware of the poison in the canteen. 65

Another popular modification, especially in criminal law, is to apply the strong necessity criterion to the competing conditions as an aggregate condition. Some, including Mackie, treat the aggregate condition as a cause while illogically denying causation by any of the individual included conditions. 66 Others illogically treat causation by the aggregate condition as establishing causation by each included condition. 67 In Germany this is known as the “formula of alternatives” (Alternativenformel): “Of several conditions which can be eliminated separately but not cumulatively without the effect failing to occur, each is a cause.” 68 Although advertised as an extension of the strong necessity criterion, this formulation instead directly contradicts it. The formula does not merely allow that the alternative cause

62. Toepel, supra note 42, at 112.
63. Id. at 120–22; accord Steel, supra note 1, at 18, 20, 20 n.23. Toepel might also take the location of the effect into account in some situations. See Toepel, supra note 42, at 121–22.
64. Toepel, supra note 42, at 120–21.
65. See, e.g., Engisch, supra note 39, at 15–16; Puppe, Der Erfolg, supra note 23, at 863, 873–74; Puppe, Concept, supra note 5, at 107; Wright, Once More, supra note 5, at 1115–20.
66. Mackie, supra note 4, at 47, apparently approved by Toepel, supra note 42, at 121; cf. Lewis, supra note 61, at 80. Steel is willing to adopt the aggregate but-for criterion for overdetermination by multiple negative conditions, but not positive ones. See Steel, supra note 1, at 32, 34.
67. E.g., Model Penal Code § 2.03 cmt. 2 at 259 (AM. LAW INST. 1985); Keeton et al., supra note 50, at 268–69.
68. This formula was first discussed in Ludwig Traeger, Der Kausalbegriff im Straf und Zivilrecht 47–48 (1904), where it was rejected because it fails in preemptive causation cases.
is not a necessary condition, it prohibits it from being a necessary condition, to avoid being able to declare every fact as a cause for any event by combining it with a fact that is really a cause of it. Furthermore, it fails to distinguish instances of duplicative causation from instances of preemptive causation (since when lumped together the preemptive and preempted conditions are strongly necessary, while neither is when considered separately) and to prevent causally irrelevant conditions from being treated as causally relevant when there is causal overdetermination (since the irrelevant conditions are not strongly necessary when considered individually, while the aggregate formed by lumping them in with the true causal conditions is strongly necessary).69

Some defenders of the strong necessity criterion have stated that, when applying it in overdetermined causation situations, you should not take into account competing conditions that prevented the condition at issue from being a strongly necessary condition.70 In Germany those using this approach refer to the competing conditions, when pre-empted, as “hypothetical,” “reserve,” or “substitute” causes.71 This approach, which operationally is the same as the aggregate strong necessity analysis discussed in the prior paragraph, similarly directly contradicts the strong necessity criterion, assumes the answer to be proved regarding duplicative versus preemptive causation, relies on the covering law analysis rather than a strong necessity analysis to reach these conclusions, and makes the outcome of the counterfactual reasoning indeterminate. If the condition at issue had not occurred, the causal process involving the competing condition would have been completed, but that now has also been removed from consideration, leaving an often very indeterminate scenario.72

In his counterfactual possible worlds theory, Lewis attempted to deal with preemptive causation situations by applying the strong necessity criterion at each step in the actual causal chain between the alleged cause and the effect at issue and treating the successful application at each step as

69. Dobbs, supra note 47, at 417; Puppe, Concept, supra note 5, at 78–79; Toepel, supra note 42, at 121; Wright, Causation, supra note 46, at 1780–81.
70. See, e.g., EUROP EAN GROUP ON TORT LAW, PRINCIPLES OF EUROPEAN TORT LAW [PETL] art. 3:102 at 46 (2005); Stapleton, Choosing, supra note 40, at 441–43 (“duplicate necessity”); Stapleton, Extended But-For, supra note 7, at 703–04 (ignoring preempted conditions). Oddly, this analysis has been adopted in PETL article 3:102 and section 27 of the Restatement Third despite its obvious defects and its failure to capture the proper weak necessity/strong sufficiency criterion that is elaborated in the comments to section 27. See infra note 113.
71. See Steel, supra note 1, at 19 (who suggests that the German courts are implicitly applying covering law analysis rather than the strong necessity criterion); Puppe, Der Erfolg, supra note 23, at 868–69, 888–89.
“causal dependence” rather than actual causation, which he states exists only if such causal dependence exists at each step, while also relying on targeted miracles to avoid backward counterfactual dependence.73 This maneuver is, again, an abandonment of the strong necessity criterion and instead relies on the covering law analysis to reconstruct the causal chain between the condition at issue and the effect. Moreover, while it will work in cases of early preemption, in which the competing condition is preempted prior to the last step in the actual causal chain, it will not work in cases of concurrent or late preemption, in which such preemption does not occur until the effect has occurred, and it will not work in duplicative causation situations.74 Lewis arbitrarily dismisses the duplicative causation situations as being “useless as test cases because [we] lack firm naïve opinions about them.”75 Although Lewis may not have had firm opinions, almost everyone else does and treats each of the duplicative conditions as a cause.

When courts have recognized situations involving causal overdetermination, they have adopted several different approaches. A few insist on compliance with the strong necessity criterion and thus erroneously deny the existence of actual causation.76 Some attempt to apply the strong necessity condition by detailing the result or aggregating the contributing conditions, as described above.77 Many recognize a condition as a cause if it was either strongly necessary or independently strongly sufficient. The independently strongly sufficient criterion encompasses situations in which a condition is necessary to complete a minimally sufficient set of actual conditions that does not contain any of the competing conditions that prevented it from being necessary for the result—for example, each of two fires when combined with the other required actual conditions (oxygen, fuel, etc.) but not the other fire.78 Unfortunately, most references to the independently

73. Lewis, supra note 4, at 563, 567; see Paul & Hall, supra note 38, at 16–17, 43–44.
74. Lewis’s counterfactual possible-worlds theory is thoroughly criticized in Moore, supra note 7, at 384–90, 392–425.
75. Lewis, Causation, supra note 4, at 567 n.12; see also Lewis, Causation as Influence, supra note 61, at 80.
76. E.g., Burrage v. United States, 134 S. Ct. 881, 887–91 (2014); cf. Price Waterhouse v. Hopkins, 490 U.S. 228, 282 (1989) (Kennedy, J., dissenting) (“Any standard less than but-for . . . represents a decision to impose liability without causation”); but see Burrage, 134 S. Ct. at 890 (majority opinion) (dicta indicating possible exceptions); supra note 53 (discussing Paroline v. United States, 134 S. Ct. 1710 (2014)).
77. See supra text at notes 54–69.
strongly sufficient criterion, including those in the first and second Restatements of Torts, fail to specify the required complete instantiation of the antecedent portion of the relevant causal generalization (and its underlying causal laws) that is required for causal sufficiency, rather than mere lawful sufficiency, and therefore fail to distinguish preempted conditions from duplicative causes.79 For example, if the victim drinks a lethal dose of poison for which there is no antidote, which under the laws of nature guarantees her death within several hours, but her head is chopped off by an axe wielder while the poison is still working its way through her system, the antecedent portion of the death by bleeding causal generalization was fully instantiated, while the antecedent portion of the death by poisoning causal generalization was not.

Other courts, primarily in common law jurisdictions, employ undefined and unelaborated phrases such as “substantial factor,” “material contribution,” “common sense causation,” or merely “contributed” or “caused” to identify as causes conditions that fail to satisfy the strong necessity criterion due to causal overdetermination.80 These phrases are of no help in resolving the causation issue but rather are merely labels applied to an unexplained conclusion. The words “factor,” “contribution,” and “causation” merely restate the causal issue. The “substantial,” “material,” and “common sense” qualifiers only make things worse by adding tests of significance that confuse the causation issue with the normative responsibility

79. Restatement of the Law of Torts § 432 (A.M. Law Inst. 1934); Restatement (Second) of the Law of Torts § 432 (A.M. Law Inst. 1965). Section 432 also erroneously treats multiple “sufficient” conditions as causes only if all of the conditions were actively operating and “sufficient.” Id.; see Wright, Once More, supra note 5, at 1097–1101. For further discussion of the critical difference between causal sufficiency and mere lawful sufficiency, see Ness Account, supra note 20, at 289, 297–303. The failure to distinguish mere lawful “guarantee” sufficiency from complete-instantiation causal sufficiency has been a major defect in Stapleton’s work. See id. at 298–99. In her most recent paper, she attempts to distinguish the two types of situations by including legally irrelevant details in the description of the relevant injury. Stapleton, Extended But-For, supra note 7, at 723–24; see supra note 54 and accompanying text.

issue. Courts in civil law jurisdictions generally employ unexplained and unelaborated findings or presumptions of causation or, usually to the same effect, reverse the burden of proof.

**B. Weak Necessity / Strong Sufficiency**

Likely as a result of their need to focus on specific instances of causation and their related exposure to recurring instances of overdetermined causation, legal theorists, rather than non-legal philosophers, have led the way in pointing out the inadequacy of the strong necessity criterion and related counterfactual theories and developing, instead, a comprehensive covering law account that encompasses the strong necessity criterion and the independently strongly sufficient criterion but goes beyond both to properly resolve the causation issue in all instances.

Karl Engisch, in a short monograph published in 1931, seems to have been the first to reject the strong necessity criterion as an exclusive criterion in favor of the more comprehensive covering law analysis. Using preemptive causation situations, he demonstrated the tautological and circular nature of attempting to detail the result to reach the correct causal conclusion when using the strong necessity criterion. The following formulation by Engisch is still often cited in the German criminal law literature:

An act—we initially only consider an actual doing—is only to be considered as being causal for a specific outcome when it is directly and through general laws of nature connected to every event following the act in a chronological order that occurred as a result of the act and the thereon following events that have materialized in any part of the facts of the case that are prohibited by law.

However, Engisch did not determine the logical form of the relation between a cause and the effect. German scholars, therefore, while praising Engisch for the theoretical superiority of his understanding of causation

---

81. Restatement (Third) of the Law of Torts: Liability for Physical and Emotional Harm § 26 cmt. j & reporters’ note (AM. LAW INST. 2010); Wright, Causation, supra note 46, at 1781–84; Wright, Once More, supra note 5, at 1075–80.
82. See Digest, supra note 47, at 256, 347–49, 457–543; Steel, supra note 1, at 19, 23; Van Dam, supra note 6, at 324–27, 329–32; Fairgrieve & G’sell-Macrez, supra note 1, at 123–24, 126–27; Spier & Haazen, supra note 1, at 146–47.
83. Engisch, supra note 39, at 14.
84. Id. at 15–16; see supra text at notes 54–65.
85. Engisch, supra note 39, at 21.
over the strong necessity criterion, generally have continued to insist upon
the strong necessity criterion as the supposed exclusive criterion.86

In 1959, Herbert Hart and Tony Honoré published their monumental
treatise, *Causation in the Law*.87 Like Engisch, they recognized that more
than one set of minimally sufficient conditions for a specific consequence
could be (and frequently is) instantiated in a specific situation.88 Relying on
and extending Mill’s covering law analysis,89 they employed the least
stringent, weak sense of necessity to identify “causally relevant factors” in
a specific situation: “A condition may be necessary just in the sense that it
is one of a set of conditions jointly sufficient for the production of the con-
sequence: it is necessary because it is required to complete this set.”90 This
can also be described as a strong sense of sufficiency, contrasted with a
strict sense that requires that the condition at issue be sufficient by itself for
the consequence (which is never true) and a weak sense that requires that it
merely be part of a sufficient set (which would allow anything, no matter
how irrelevant, to be treated as a cause by adding it to an already sufficient
set).91

Hart and Honoré’s weak necessity definition of a “causally relevant
factor” was a major advance in the analysis of causation, in both law and
philosophy.92 However, the potential clarifying impact of their analysis was
seriously blunted by several aspects of their discussion. Most significantly,
their analysis of causally relevant factors was only a preliminary step, not
even mentioned in the introduction to their book, in pursuit of their primary
project, which was an unsuccessful attempt to use ordinary language analy-
sis to identify supposedly factual, non-normative causal criteria, applicable
in non-legal as well as legal contexts, for treating only one or a few of the
many causally relevant factors in a specific situation as causes.93 This pri-

86. Noted, critically, by Puppe, *Concept*, supra note 5, at 69–70, 70 n.3. See infra text at notes
95–96.
87. H.L.A. Hart & A.M. Honoré, *Causation in the Law* (1959); see also Tony Honoré, *Die
Kausalitätslehre im anglo-amerikanischen Recht im Vergleich zum deutschen Recht*, 69 ZStW 465
(1957).
89. *Id.* at 15–22, 111–13. Although they were aware of Engisch’s prior work, they only men-
tioned it in a footnote toward the end of their treatise. *Id.* at 435 n.18 (1st ed. 384 n.4).
90. *Id.* at 112–13.
91. Wright, *Pruning*, supra note 32, at 1020; see Mackie, *supra* note 4, at 38–40; Puppe, *Der
Erfolg*, supra note 23, at 865–69, 875–76.
92. See, e.g., Philippa Foot, *Hart and Honoré: Causation in the Law*, 72 Phil. Rev. 505–15
(1963); M.P. Golding, *Causation in the Law*, 59 J. Phil. 85 (1962); D.D. Raphael, *Causation in
the Law*, 37 Phil. 83 (1962).
93. Hart & Honoré, supra note 4, at 1–6, 11–13, 23–108, which is criticized in Haskell Fain,
*Hart and Honoré on Causation in the Law*, 9 Inquiry 322 (1966); Richard W. Wright, *The Nightmare
and the Noble Dream: Hart and Honoré on Causation and Responsibility*, in *The Legacy of H.L.A.*
mary focus overwhelmed and sometimes distorted their analysis of causally relevant factors, which received minimal notice in the legal literature, and instead promoted further confusion by courts, lawyers, and philosophers of the distinct issues of causation and responsibility.\(^{94}\)

In the early 1980s, we each began writing (in different languages and with different doctrinal backgrounds and thus until recently unaware of each other’s work) a series of papers criticizing the strong necessity criterion and instead insisting on the covering law theory’s weak necessity/strong sufficiency criterion. In 1980, Ingeborg Puppe, referencing Engisch’s work but being unaware of Hart and Honoré’s and Mackie’s work, drew on Hempel and Oppenheim’s deductive-nomological theory of causal explanations in the philosophy of science\(^{95}\) to supply what Engisch had not: a specification of the proper logical form of a causal condition, which is weak necessity/strong sufficiency. As Puppe often puts it, a cause is a necessary element of a true causal explanation of the result according to causal laws.\(^{96}\) In 1985, Richard Wright sought to revive, clarify, correct and extend Hart and Honoré’s weak necessity analysis of causally relevant factors, which he called “NESS” (necessary element of a sufficient set) conditions, while rejecting Hart and Honoré’s attempt to treat only some of the causally relevant factors as causes.\(^{97}\)

The weak necessity criterion underlies—and is relied upon when employing—both the strong necessity criterion and the independently strongly sufficient criterion. Each is merely a corollary of the weak necessity criterion that works only in certain situations. The strong necessity criterion works only if there was a single minimally sufficient set of actual conditions in the specific situation, in which case every condition that was necessary for the complete instantiation of that set was also necessary for the result. The independently strongly sufficient criterion encompasses situations in which a condition is necessary to complete a minimally sufficient set that does not contain any of the competing conditions that prevented it from being strongly necessary. Both criterions are only special cases of weak necessity/strong sufficiency. When there is duplicative overdetermi-
nation by independently strongly sufficient conditions (e.g., two or more “independently sufficient” fires), you can formulate two or more minimally sufficient sets of conditions, each of which contains only one of the competing conditions (fires) in addition to the other required conditions, some of which (e.g., oxygen, dry fuel, and an existing property to burn down) will be common to each minimally sufficient set.98

Going beyond prior discussions of the weak necessity (strong sufficiency) criterion, which only encompassed conditions that were strongly necessary or independently strongly sufficient, we each have noted that the weak necessity criterion is able to recognize as causes conditions that were neither strongly necessary nor independently strongly sufficient but were strongly sufficient when included in a minimally sufficient set containing one or more other conditions of the same type. Consider, for example, the casting of three individual affirmative votes when only two were necessary for approval of some action, or three sources of force, fire, water, noise or other type of condition, each of size X, which combine to produce an indivisible injury for the occurrence of which only 2X amount was necessary. Employing the weak necessity analysis, each of the affirmative votes or sources can properly be identified as a cause of the relevant legal injury by including it and either one of the other two affirmative votes or sources in a completely instantiated minimally sufficient set that does not include the third affirmative vote or source.99

Wright argues that the weak necessity analysis can and should be employed to recognize even more conditions as causes. For example, assume there were only two voters, X with the right to cast 2N votes and Y with the right to cast only N votes, who each cast their votes in the affirmative when 2N affirmative votes were necessary and sufficient for approval of an action detrimental to the plaintiff, or two simultaneously operative sources of force, fire, water, noise, toxin, etc., X of size 2N and Y of size N, when amount 2N was necessary and sufficient for the plaintiff’s suffering of an indivisible injury. In each case, X’s contribution was both strongly necessary and independently strongly sufficient and thus clearly was a cause. Y’s

98. See supra text at notes 78–79.

99. See Ingeborg Puppe, Entscheidungsanmerkung zu BGHSt 37,107 [1992] JURISTISCHE RUNDSCHAU 30, 32; Puppe, Concept, supra note 5, at 98; Wright, Causation, supra note 46, at 1791–93. Honoré has noted, without any explanation, that such conditions could be causes. A.M. Honoré, Causation and Remoteness of Damage, in 11 INTERNATIONAL ENCYCLOPEDIA OF COMPARATIVE LAW, TORTS 7-107, 7-108, 7-115, 7-121 (André Tunc ed., 1983). For examples of such situations, which are not rare, see, e.g., Bundesgerichtshof [BGH] [Federal Court of Justice] 1990, NEUE JURISTISCHE WOCHENSCHRIFT [NJW] 2560, 2566 (Ger.) (votes); Warren v. Parkhurst, 92 N.Y.S. 725 (Sup. Ct. 1904), aff’d, 93 N.Y.S. 1009 (App. Div. 1905), aff’d, 78 N.E. 579 (N.Y. 1906) (pollution); DIGEST, supra note 47, at 531–43; STEEL, supra note 1, at 21–23; infra note 115 and accompanying text.
contribution, although neither strongly necessary, independently strongly sufficient, nor, apparently, even strongly sufficient (necessary to complete a minimally sufficient set which may include other conditions of the same type), can be shown to be strongly sufficient (and thus a cause) by including it in a minimally sufficient set that also includes X’s contribution described as “at least size N.” When compared to the situations discussed in the prior paragraph, it cannot possibly matter, from the perspective of scientific/natural/actual causation, which person or source supplied the various votes or amounts of force, fire, water, noise, toxin, etc.\(^{100}\)

Puppe disagrees. She insists on applying the weak necessity criterion to each actual condition in a specific situation and is unwilling to apply the “at least” descriptive technique to X’s conduct or activity to enable treating Y’s conduct or activity as a cause. Instead, she argues that the “at least” descriptive technique should be used to describe the relevant abstract condition in the causal law to enable (i) (similar to Mackie) X’s conduct or activity to be treated as an instantiation of that abstract condition even if it was quantitatively greater than the amount necessary for such instantiation, (ii) multiple actual conditions each of which is by itself insufficient for such instantiation to be treated as causes, but not to allow (iii) any actual condition which was not by itself sufficient for such instantiation to be treated as a cause when there is another actual condition that by itself was sufficient.\(^{101}\)

The use of subsets and/or “at least” descriptions of actual conditions to make up a minimally sufficient set of actual conditions in a specific instance has been challenged by some as supposedly assuming that the omitted conditions or their more-than-at-least properties did not exist. This is not true; rather, they simply are not included in the actual conditions chosen to make up the specified minimally sufficient set. Their actual existence outside of the specified set must be considered to make sure that they did not prevent the complete instantiation of the specified set by preventing the instantiation of one or more of its required elements.\(^{102}\)

\(^{100}\) Wright, Causation, supra note 46, at 1793–94; Wright, Pruning, supra note 32, at 1035–39. If Y’s contribution is trivial in comparison to X’s, liability generally will be denied due to a non-causal normative limitation on attributable responsibility. Restatement (Third) of the Law of Torts: Liability for Physical and Emotional Harm \(\S\) 36 (Am. Law Inst. 2010); Wright, Legal Responsibility, supra note 6, at 1448–50, 1450 n.84.

\(^{101}\) See Mackie, supra note 4, at 43–44, 65, 153, 265; Puppe, Der Elforg, supra note 23, at 883–95.

\(^{102}\) Wright, Causation, supra note 46, at 1793–94; Wright, NESS Account, supra note 20, at 304; see Puppe, Concept, supra note 5, at 93–94, 98.
Nevertheless, especially when “at least” descriptions of some conditions are used to allow other conditions to be recognized as causes, the weak necessity criterion may be too conceptually complex to be applied directly to singular instances in actual practice.103 This difficulty is most apparent in the frequent situations in which multiple sources of unknown size combine to cause some effect, as in the two cases most often cited (erroneously) in the United States as supposed examples of causation by multiple independently sufficient conditions, Anderson v. Minneapolis, St. Paul & Sault Ste. Marie Railway Co.104 and Corey v. Havener.105 In these two cases, the courts did not require proof—and it is doubtful that it could have been proved—that the defendant’s tortious conduct was strongly necessary, independently strongly sufficient, or even strongly sufficient, but rather merely that it was a “substantial factor” (Anderson) or “contributed” to the injury (Corey).106 Similarly, it often will be difficult to prove, even using the weak necessity criterion with the subset or “at least” descriptive technique, that one reason among many for some decision or action was a cause.107

Wright argues that the conceptual complexities can be avoided and the informational difficulties greatly reduced by applying the weak necessity criterion only to the construction of causal laws as minimally sufficient sets of abstract conditions, while treating any actual condition that was a coherent part of the complete instantiation of the antecedent portion of a relevant causal law in a specific situation as a cause of the instantiated consequent portion, even if the abstract condition that it helped to instantiate was over-instantiated.108 To be a coherent part of the complete instantiation it must


107. See supra note 103.

be consistent with the other required conditions. For example, a second fire that arrived after the house had already burnt down would not be consistent with the requirement that there be a house to burn down when the fire arrives at the location of the house. This less stringent way of understanding and applying the weak necessity criterion in overdetermined causation situations is consistent in result with the direct application of the criterion to the actual conditions in the specific situation using “at least” descriptions of some of those conditions. However, it is more transparent and simple to use, while also enabling one to reach the correct causal conclusion in cases like Anderson and Corey in which it cannot be proved that a condition was strongly necessary or strongly sufficient but it nevertheless clearly contributed. Similarly, to establish that some information contributed to a specific decision, it need only be established that the information was considered by the subject and counted positively in favor of the decision; if so, it was part of the complete instantiation of a minimally sufficient set of reasons for that decision.109

Puppe continues to insist that the weak necessity criterion be applied directly to the actual conditions in the specific situation. She suggests that, in cases where we cannot figure out the exact amount of a minimal sufficient condition for an effect and therefore have to work with “at least” descriptions, the “at least” descriptions should be put into the causal law to avoid distinguishing, in our application of the weak necessity criterion,
between the causal law and its application in a single case and thereby giving up the direct application of the logical form of the weak necessity criterion as a requirement for a condition to be a cause in a single case. She acknowledges that this approach is imprecise since we do not thereby get a minimally sufficient set of actual conditions in the single case, and that we must be aware of this in cases where different actors contribute to the amount for which we only give an “at least” description. However, she notes, if only one person provides this amount, that would not cause a failure, because the minimal sufficient amount is contained in the “at least” description. Wright observes that this approach will not help in cases like Anderson and Corey, for which we do not know the actual or relative sizes of the distinct contributions.

The weak necessity criterion, including its application to conditions that were neither strongly necessary nor independently strongly sufficient, has been generally accepted as a major improvement over the strong necessity criterion by legal academics in common law jurisdictions\(^\text{110}\) and increasingly in other jurisdictions as it has become better known,\(^\text{111}\) although some have questioned Wright’s claim to have captured, non-circularly, the essence of the concept of causation and/or its application in some situations, especially overdetermined negative causation situations.\(^\text{112}\) It has

\(^{110}\) See, e.g., Steel, supra note 1, at 25–33; Schaffer, Contrastive Causation in the Law, supra note 36, at 285; Stapleton, Causation in the Law, supra note 3, at 765–67; Gemma Turton, Using NESS to Overcome the Confusion Created by the 'Material Contribution to Harm' Test for Causation in Negligence, 30 J. PROF. NEC. L. 50 (2014); Euan West, The Utility of the NESS Test of Factual Causation in Scots Law, 4 ABERDEEN STUDENT L. REV. 39, 47 (2013); authors cited in Wright, NESS Account, supra note 20, at 285 n.1. Those who claim to be applying some modification or extension of the strong necessity criterion rely, explicitly or implicitly, on the weak necessity covering law account. See, e.g., Green, supra note 103, at 13–14 (who, as Steel, supra note 1, at 34 n.81 has noted, needs to rely on the NESS weak necessity criterion or some similar criterion to determine whether a condition was “operative”); Stapleton, Choosing, supra note 40, at 474 (stating that the NESS weak necessity criterion best elaborates her concept of “involvement,” including “necessity” and “duplicate necessity” as well as “contribution”). In her most recent paper, Stapleton defines a condition as a cause of some phenomenon “only if, but for that factor alone, (i) the phenomenon would not exist or (ii) an actual contribution to an element of the positive requirements for the existence of the phenomenon would not exist”). Stapleton, Extended But-For, supra note 7, at 713. Part (ii) is viciously circular, incomplete and murky. It fails to define “contribution” or to require instantiation of all “the requirements for the existence of the phenomenon,” and her discussion of its application to deny causation by negative conditions that are not strongly necessary is not forensically friendly, persuasive or consistent. See id. at 714–23; Steel, supra note 1, at 25, 35–36.


\(^{112}\) The circularity objection is discussed and rejected in Wright, NESS Account, supra note 20, at 288–90. Our previously differing analyses of overdetermined negative causation situations are recon-
been adopted in the Restatement (Third) of Torts as the correct analysis of causation, but the black letter sections fail to capture it accurately and thus continue to be unable to distinguish between duplicative and preemptive causation situations, and the comments, which rely upon it, fail to emphasize the required complete instantiation sense of sufficiency. Although the courts generally continue to resort to conclusory, unelaborated, and unexplained findings of causation when the strong necessity criterion erroneously fails to identify causation, often using phrases such as “substantial factor” or “material contribution” or unexplained presumptions, few as yet have been made aware of the weak necessity criterion. When aware of and properly understanding it, the courts have generally accepted or at least relied upon it.

V. PROVING CAUSATION

Proof of a singular instance of causation requires proof of (1) a scientifically valid causal law or generalization (the abstract “general causation” or causal capacity issue), and (2) complete instantiation of the allegedly relevant causal generalization and its underlying causal laws in the specific situation (the concrete “specific causation” issue). General causation is usually assumed, without requiring any proof, in traditional crash-bash-
slash situations, but it has become a subject of major contention in modern
tort cases involving, for example, toxic or carcinogenic substances, in
which the relevant causal laws often are much less well understood and
even doubted as being deterministic.\textsuperscript{116}

A significant statistical correlation between the occurrences of two
different conditions is an indication, but never by itself sufficient proof,
that they are connected as abstract elements in a causal law. It may instead
be a spurious correlation between the nodes of an epiphenomenal relation,
each due to a common cause, or an accidental, non-causally related correla-
tion. Before concluding that a general causal relation exists between condi-
tions of type $A$ and $B$, scientists generally require, in addition to a strong,
consistent, specific, coherent, graded-to-exposure, testable statistical asso-
ciation, that $A$ occur before $B$, that there be some scientifically plausible
explanation or hypothesis regarding how $A$ causes $B$, and that other possi-
bile explanations for the correlation be ruled out.\textsuperscript{117}

As Mill emphasized, we rarely, if ever, will have complete knowledge
of the relevant causal laws and actual conditions in a specific situation.
Instead, we employ causal generalizations, which lump together at a macro
level, with very incomplete specification, a large number of underlying
ciausal laws at the micro level. Since a causal generalization is not a com-
plete specification of the underlying causal laws, proof of instantiation of
even all of the antecedent conditions in the causal generalization provides
only an aggregate class-based statistical probability that the underlying
ciausal laws were completely instantiated and, thus, that the condition at
issue actually was a cause of the relevant consequence. Nevertheless, the
proven actual conditions in the specific situation, when considered in light
of alternative causal stories, may provide sufficient direct and circumstan-
tial evidence regarding instantiation of the network of causal processes in
the specific situation to warrant the formation of a belief, rather than a mere
class-based statistical probability, that a specific causal generalization and
its underlying causal laws were instantiated in the specific situation.\textsuperscript{118}

\textsuperscript{116} See \textit{Restatement (Third) of the Law of Torts: Liability for Physical and
Emotional Harm} § 28 cmt. c (Am. Law Inst. 2010); Steel, supra note 1, at 76.

\textsuperscript{117} See \textit{Susan Haack, Evidence Matters: Science, Proof, and Truth in the Law} 239–63
(2014); Steel, supra note 1, at 68–75; Austin Bradford Hill, \textit{The Environment and Disease: Associa-
tion or Causation?}, 58 Proc. Royal Soc. Med. 295 (1965); Symposium on Legal and Scientific Per-

\textsuperscript{118} See Haack, supra note 117, at 17–19, 52–56, 60–61; Wright, \textit{Pruning}, supra note 32, at
1044–54; Richard W. Wright, \textit{Proving Causation: Probability versus Belief}, in \textit{Perspectives on
Causation}, supra note 1, 195, 205 [hereinafter Wright, \textit{Proving Causation}]; supra notes 29–30 and
accompanying text.
The required warranted degree of belief in criminal cases is now generally acknowledged to be “beyond a reasonable doubt” in common law jurisdictions and the same or “a virtual certainty” in civil law jurisdictions. There is much less agreement on the required warranted degree of belief in civil cases. In all common law jurisdictions, some civil law jurisdictions (e.g., China, Italy, Switzerland, and Thailand), and “mixed” (e.g., Scandinavian) jurisdictions, the standard of persuasion in civil litigation is stated to be much lower than in criminal litigation: at best, a bare minimal belief, despite perhaps substantial doubts, described as a “preponderance of the evidence” in the United States and some civil law jurisdictions and as a “balance of probability” in the British Commonwealth and Scandinavia.\(^\text{119}\)

In many civil law jurisdictions, the academic doctrine considers the required warranted degree of belief in civil cases to be the same high standard as in criminal cases, although the civil codes generally merely require

\(^{119}\) See, e.g., Y v. Norway, 2003-II Eur. Ct. H.R. 161, 163, http://echr.coe.int/Documents/Reports_Recueil_2003-II.pdf (affirming civil liability under Norway’s “balance of probability” standard after acquittal under the “reasonable doubt” standard for criminal liability); OBLIGATIONENRECHT, CODE DES OBLIGATIONS, CODICE DELLE OBLIGAZIONI [Code of Obligations] Jan. 1, 2014, RO 220, art. 53 (Switz.) (“When determining fault or lack of fault and capacity or incapacity to consent, the court is not bound by the provisions governing criminal capacity nor by any acquittal in the criminal court. The civil court is likewise not bound by the verdict in the criminal court when determining fault and assessing compensation.”); Wright, Proving Causation, supra note 118, at 195–96. In Thailand, sentence two of article 227 of the Criminal Procedure Code states: “Where any reasonable doubt exists as to whether or not the accused has committed the offence, the benefit of doubt shall be given to him.” CRIM.P ROC.C ODE § 227 [p. 87] (Thail.).

In contrast, article 104 of the Civil Procedure Code of Thailand states, “The Court shall have full power to decide whether the evidence as adduced by the parties is relevant to the issue and is sufficient to be taken as conclusive or not and then to give judgment accordingly.” CIV. PROC. CODE § 104 [p. 55] (Thail.), http://www.imolin.org/doc/amld/Thailand_Criminal%20Procedure%20Code.pdf. The latter provision has been interpreted as requiring the court to decide the case by comparing the weight of the evidence of both parties. PANYA SUTHIBODEE, THE LAW OF EVIDENCE 11 (Ramkhamhaeng Univ. Press 1998). The situation is similar in the People’s Republic of China. Compare CRIMINAL PROCEDURE LAW OF THE PEOPLE’S REPUBLIC OF CHINA (promulgated by the 11th National People’s Conference, effective 2012) art. 53, http://www.inchinalaw.com/wp-content/uploads/2013/09/PRC-Criminal-Procedure-Law-2012.pdf (“Evidence shall be deemed to be sufficient and concrete” when “the ascertained facts have been proved beyond reasonable doubt”), with PROVISIONS OF THE SUPREME PEOPLE’S COURT ON EVIDENCE IN CIVIL PROCEEDINGS (Promulgated by the Supreme People’s Court, effective April 1, 2002) art. 73, http://seafarersright.org/wp/wp-content/uploads/2014/11/CHN_LEGISLATION_EVIDENCE-IN-CIVIL-PROCEDURES_2001_ENG.pdf (“Where both parties concerned respectively produce contrary evidence to prove the same fact, but neither party has sufficient grounds to negate each other’s evidence, the people’s court shall, in light of the case details, decide whether the probative force of one party’s evidence is obviously more powerful than that of the evidence from the other party, and confirm the evidence with the more powerful probative force.”). However, in Thailand and a number of other civil law jurisdictions, it is (illogically) assumed that a failure to prove certain facts under the higher criminal law standard of persuasion conclusively establishes the non-existence of those facts under the lower civil litigation standard. See, e.g., Thailand Supreme Court decisions 349/2555, 1144/2553, and 9209/2553, interpreting CODE OF CRIMINAL PROCEDURE art. 46, which states “In giving judgment in the civil claims, the Court shall be bound by the facts as found by the judgment in the criminal claims”.)
that the judges be convinced of the truth of the matter at issue, without specifying any particular degree of belief, and there is reason to believe that in practice the required degree of belief is lower in civil cases.120

A literal interpretation of the “balance of probability” standard and a similar common interpretation of the “preponderance of the evidence” standard as merely requiring a “more likely than not” probability (rather than belief) have led many academics and some courts in common law jurisdictions to assume that all that is required for proof of causation in a specific instance is an aggregate fifty-plus percent class-based statistical probability.121 As courts in Australia and civil law jurisdictions generally state, courts in common law jurisdictions often state, and the British Supreme Court has recently affirmed, the standard of persuasion in both criminal and civil law cases requires formation of an actual belief, rather than reliance on a mere statistical probability, to prove that causation actually existed in a particular situation.122 An abstract class-based statistical probability of causation, no matter how high, does not instantiate any element in any causal law or generalization, nor does it otherwise provide any help in determining whether the causal laws underlying the generalization were fully instantiated in the specific situation. Only concrete evidence of the actual conditions in a specific situation is capable of participating in such instantiation.123

Puppe agrees that mere statistical probability is not sufficient proof of actual causation for deterministic causal processes, but she believes that certain processes, especially the development of cancer, are partly indeterministic, and she has offered a method to adapt the NESS covering law
account for these processes: instead of a minimal sufficient condition we have to formulate a complete description of the objective probability. A condition is a legitimate element of the lawful explanation of an objective probability if it increases this probability, and if this is the case the one who caused this condition is responsible for the actual result.\textsuperscript{124} Wright disagrees for several reasons, including, especially, the fact that this approach apparently would treat any condition that merely increases the risk, \textit{ex ante}, as a cause of the actual harm.

Courts today increasingly have to deal with situations in which specific causation cannot be proven or disproven, due to insufficient knowledge of the relevant causal laws and/or the actual conditions in the specific situation or to probabilistic elements in the relevant processes. Some courts in common law jurisdictions and even a few in civil law jurisdictions, relying on probabilistic interpretations of the applicable “balance of probability” and “preponderance of the evidence” standards of persuasion, have treated proof of a mere class-based statistical probability of causation as sufficient proof of actual causation in some situations, most notably in cases involving toxic or carcinogenic substances or medical failures to properly diagnose or treat an adverse medical condition. In other situations, for example, those involving theoretically divisible but practically indivisible injuries or practically indistinguishable defendants, they have recognized the fallacy of doing this. Like some courts in civil law jurisdictions, they instead have considered and sometimes adopted second-best liability provisions by employing presumptions, reversing the burden of proof, or imposing proportional liability based on the probability of causation.\textsuperscript{125}

The second approach, which recognizes the impossibility of proving or disproving specific causation in these situations, is the one that should always be followed.\textsuperscript{126} Failure to do so and instead treating mere statistical probability as sufficient proof of actual causation leads to numerous paradoxes and descriptively and normatively unsustainable results.\textsuperscript{127} When proof of specific causation or its lack is impossible due to insufficient knowledge of the relevant causal laws/generalizations and/or the facts in a

\begin{thebibliography}{99}
\bibitem{126} See \textit{Fairgrieve & G’sell-Macrez, supra} note 1, at 123–24, 126–27; Moréteau, \textit{supra} note 47, at 792–801; Wright, \textit{Proving Causation, supra} note 118, at 205–20.
\end{thebibliography}
specific situation and all we have is an aggregate statistical probability of specific causation, this should be acknowledged as well as the normative issue of appropriate legal responsibility given such irresolvable uncertainty.

VI. CAUSATION IN EFFICIENCY THEORY

The efficiency theorists assume that the purpose of tort law and all other areas of law is or should be the maximization of aggregate social welfare, with welfare defined (inconsistently, often in the same paper) as utility (pleasure or preference satisfaction, in which case efficiency theory is simply utilitarianism), wealth as measured by persons’ willingness and ability to pay to obtain some resource that they do not have (rather than their usually different willingness to sell it if they already have it), or market value, all of which are quite different normatively and operationally.128 Although the efficiency theorists agree on the basic goal (maximization of aggregate social welfare, however defined or measured), they disagree on whether and how tort law serves this goal.

Guido Calabresi, who along with Ronald Coase initiated explicit efficiency analyses of legal responsibility,129 broke the basic maximization of social welfare goal down into three subgoals: efficient deterrence (minimization, ex ante, of expected “primary accident costs,” which are the harms suffered by individuals as the result of risk-creating conduct and the precaution costs required to reduce the relevant risks), efficient compensation (minimization, ex post, of “secondary accident costs,” which are the disutilities suffered by the victims of actual harm, by shifting those losses to wealthier individuals (“deep pockets”) or spreading them as widely as possible), and efficient administration (minimization of “tertiary accident costs,” which are the costs to all involved of processing tort liability claims).130 Since efficient compensation conflicts with efficient deterrence and is best accomplished by social insurance, with tort liability being one of the worst possible alternatives, post-Calabresi efficiency theories of tort law have ignored efficient compensation and instead focused on efficient

deterrence and efficient administration (minimizing the sum of primary and tertiary accident costs).131

Given their focus on efficient deterrence, the efficiency theorists have had a hard time explaining any aspect of tort law, despite its being the focus of all of their initial writings on legal responsibility.132 The biggest problem for the efficiency theorists is the causation requirement, with its ex post focus on past injuries and their causes rather than ex ante focus on future expected harm. Initially, the efficiency theorists tried to minimize the practical importance of the causation requirement. They argued that, since injuries always result from the interaction of the plaintiff’s and defendant’s activities, each activity is a cause of the particular interaction and its effects, so the causation requirement is trivially satisfied and the only significant issue is how liability should be assigned among the interacting activities to minimize aggregate social costs or, equivalently, maximize aggregate social benefits.133

However, this argument does not explain or justify tort law or its causation requirement. It does not explain why liability is limited to those situations in which an injury has occurred and is further limited to harm that was caused by the tortious aspect of the defendant’s conduct or, even if liability is so limited, why the plaintiff should obtain compensation for the injury.134 As Calabresi has explained,135 since efficient deterrence focuses on providing ex ante incentives for efficient behavior rather than ex post compensation of injuries, liability equal to the expected harm could be


133. E.g., John Prather Brown, Toward an Economic Theory of Liability, 2 J. LEGAL STUD. 323, 326 (1973); Calabresi, supra note 129, at 505–06, 506 n.24; Coase, supra note 129, at 2, 13, 19, 27.

134. Wright, Bane, supra note 132, at 438.

imposed on the actor whenever his conduct creates risks to others (to force him to internalize the expected costs to others), or whenever he behaves inefficiently, or—to reduce enforcement costs—at least whenever his conduct or activity contributes to an injury (with a multiplier to adjust for the frequency of proven actual injuries compared to expected injuries), regardless of whether the injury was caused by the inefficient aspect of his conduct or activity. Such liability could be imposed through fines payable to the state, rather than providing any compensation to the person, if any, suffering an injury.136

Subsequently, acknowledging the historical persistence and normative significance of causation as a requirement for tort liability, Calabresi attempted to provide an efficiency explanation for the requirement. To facilitate doing so, he initiated a semantic maneuver that has since been copied by other efficiency theorists. Relying on and expanding the confusion between causation in its basic sense and so-called “proximate” or “legal” causation, he treated causation as a linguistically manipulable term to be employed as desired to achieve “certain [aggregate social welfare] goals that have come to be accepted as crucial to the law of torts.”137 He treated as causal concepts not only the usual strong necessity (sine qua non) criterion for actual causation, but also the attributable responsibility limitations on liability for tortiously caused injury, traditionally and misleadingly referred to as “proximate causation.” He also included a “causal linkage” concept, which is more accurately described as a “probabilistic linkage,” since it merely refers to ex ante increased risk: an act is “causally linked” to an actual injury or even a merely potential injury, whether or not it contributed to the actual injury or an injury even occurred, if the recurrence of the act in the future will increase the risk of occurrence of a similar injury.138

136. Posner argues that full compensation must be provided to persons injured by defendants’ inefficient conduct to provide an incentive for such victims to sue and to prevent such victims from engaging in inefficient excessive precaution against potential injury. RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW § 6.10 (8th ed. 2011). However, if a fine payable to the government is properly set to obtain efficient behavior by potential injurers, the victim who is not expecting compensation will minimize her expected injury and precaution costs by engaging in efficient rather than excessive precaution. See Brown, supra note 133, at 340–42; Richard W. Wright, Allocating Liability Among Multiple Responsible Causes: A Principled Defense of Joint and Several Liability for Actual Harm and Risk Exposure, 21 U.C. DAVIS L. REV. 1141, 1170–74 (1985). There is no need to limit enforcement to victims rather than whoever might be the most efficient enforcer of the liability rules (e.g., trial lawyers) or to turn over to the enforcer all of the fine rather than only a portion of it as a sufficient (contingent fee) incentive for enforcement action.


138. Id. at 71–72; see Wright, Bane, supra note 132, at 439–40, 440 n.47.
2016] CAUSATION 497

When Calabresi evaluated the usefulness of these three concepts in promoting the minimization of aggregate social costs, he concluded that none of them were useful and indeed all were detrimental for achieving efficient compensation.\(^\text{139}\) He also concluded that, while the “causal linkage” concept is highly relevant for both “specific deterrence” (prevention of activities deemed to be inefficient per se or otherwise undesirable) and “market deterrence” (deterrence through pricing mechanisms, including liability costs, of inefficient behavior), the only one of the concepts that addresses actual causation (the “but for” test) is inconsistent with “specific deterrence”\(^\text{140}\) and is only one of several alternatives, but not the best or even close to adequate, for achieving “market deterrence,” by providing data for the creation of “an actuarial basis from which actors can decide whether future safety costs are cheaper or more expensive than future injury costs.”\(^\text{141}\) He thus attempted to recast tort law’s actual causation requirement as a “flexible and functional” term that does not refer to any “inherent, ‘natural’ relationships” but instead, because of its “historical, common law gloss” and moral appeal, can and should be used to describe efforts to “identify those pressure points that are most amenable to the social goals we wish to accomplish.”\(^\text{142}\)

Other prominent efficiency theorists have followed Calabresi’s lead in attempting to recast the \textit{ex post} actual causation analysis as an \textit{ex ante} efficiency analysis. The first to do so was Steven Shavell, who employs the term “probabilistic cause” rather than “causal linkage” to refer to mere increased risk and, like Calabresi, includes it as a supposed “basic notion of causation.”\(^\text{143}\) Using unrealistic and inconsistent assumptions (including courts’ having perfect information), contrived illustrations, and defective arguments and mathematical formulas,\(^\text{144}\) he claims to prove that tort law’s actual causation requirement, which, like Calabresi, he equates with “but for” causation, is consistent with and indeed required for efficient deterrence under a strict liability regime.\(^\text{145}\) (He admits that under a negligence

139. Calabresi, supra note 137, at 73–77; see Wright, Bane, supra note 132, at 440.
140. Calabresi, supra note 137, at 79; see CALABRESI, supra note 130, at 269–70, 370 n.5; Wright, Bane, supra note 132, at 440–41.
141. Calabresi, supra note 137, at 85; see id. at 85–86; CALABRESI, supra note 130, at 251, 251 n.8, 247–49, 257–58, 267 n.2; Wright, Bane, supra note 132, at 441–42.
142. Calabresi, supra note 137, at 106–08.
liability regime the causation requirement undermines rather than promotes efficient deterrence.\textsuperscript{146} However, his arguments and supposed proofs do not employ the \textit{ex post} inquiry into what actually happened, irrespective of what was foreseen or known beforehand, that is the essence of the factual causation inquiry. Rather, he relies on supposed perfect knowledge by actors, prior to engaging in some conduct or activity, of the subsequent play-out of the risks created by that conduct or activity—that is, on \textit{ex ante} “probabilistic causation” rather than \textit{ex post} actual causation.\textsuperscript{147} Moreover, his arguments and proofs involve situations in which there are multiple actual (duplicative) or hypothetical (preempted) independently sufficient causes. In such situations, contrary to his assertions, the courts usually recognize that actual causation exists despite a lack of but-for causation, with liability depending on whether the “no worse off” limitation on attributable responsibility applies.\textsuperscript{148}

In their article on causation, William Landes and Richard Posner repeat Shavell’s error of equating \textit{ex ante} increased risk with \textit{ex post} actual causation, but with an opposite effect. Shavell uses \textit{ex ante} risk analysis and claims he is using \textit{ex post} causal analysis, while Landes and Posner use \textit{ex post} causal analysis and claim they are using \textit{ex ante} risk analysis. This is most evident in their discussion of \textit{Weeks v. McNulty},\textsuperscript{149} a case in which the defendant, McNulty, failed to install the statutorily required fire escapes in his hotel. The victim, Weeks, died in the hotel as a result of a fire. The court concluded that McNulty was not liable, despite the negligent failure to install the fire escapes, since the evidence proved that Weeks would not have tried to use the fire escapes even if they had been installed: an actual causation rationale. Landes and Posner argue:

[T]o reason thus is to assume the conclusion. We want a ground of decision that will not depend explicitly on any notion of cause and we find it in the [Hand formula’s aggregate risk-utility analysis of negligent conduct]. Evidence that the fire escapes would not have averted Weeks’s death means, in the context of our model, that the probability of his death

\textsuperscript{146.} \textit{Id.} at 485–86. Shavell argues that when courts do not have perfect information “a significant element of strict liability is inherent in the negligence rule” and thus that “all the conclusions reached about the scope of liability under strict liability [would be] relevant under the negligence rule.” \textit{Id.} at 489. However, the existence of imperfect information does not make negligence liability equivalent or even similar to strict liability. See Grady, \textit{supra} note 144, at 118–21; Wright, \textit{Bane, supra} note 132, at 451–52; \textit{infra} notes 154–57 and accompanying text.

\textsuperscript{147.} Wright, \textit{Bane, supra} note 132, at 445–46.

\textsuperscript{148.} \textit{Id.} at 446–48; see Wright, \textit{Legal Responsibility, supra} note 1, at 1434–52.

\textsuperscript{149.} \textit{Weeks v. McNulty,} 48 S.W. 809, 810 (Tenn. 1898).
was independent of whether or not a violation occurred... and hence that due care with respect to [installing fire escapes] was zero.  

This is an *ex post* causal analysis masquerading as an *ex ante* risk analysis. The failure to install the fire escapes unreasonably increased the *ex ante* risk of injury to every guest in the hotel, including Weeks, and thus was negligent (as the legislature determined), although as it turned out it did not contribute to Weeks’s death. Landes and Posner similarly employ *ex post* causal analysis while claiming to be employing the Hand formula’s *ex ante* analysis of negligence in their discussions of other types of situations. If they consistently employed *ex post* knowledge of what actually occurred to modify (replace) the analysis of *ex ante* risk, they would end up imposing negligence liability in many cases in which the defendant’s behavior was not negligent but nevertheless caused an injury. For example, assume an automobile driver is involved in a serious accident despite careful driving. Using their method of taking into account what actually occurred, as revealed *ex post*, to calculate the *ex ante* required care, the “expected” cost of failing to exercise additional care was a 100 percent probability of a serious accident, which would almost always require much higher care than the driver exercised, including, if nothing else would prevent the accident, forgoing driving that day.

The attempts by Calabresi, Landes, Posner, and Shavell to replace the actual causation requirement with an *ex ante* or *ex post* analysis of negligent (presumably inefficient) conduct continue to be accepted and repeated by efficiency theorists, despite the obvious serious defects in these attempts.


151. *Id.* at 119–22; see Wright, Bane, *supra* note 132, at 453–55; Wright, *Legal Responsibility, supra* note 1, at 1520–21, 1521 n.320.

152. Landes and Posner responded to this criticism by continuing to insist that they are using *ex ante* risk analysis rather than *ex post* causal analysis:

[Wright’s] criticism is incorrect. The point is that the owner’s carelessness [in the *Weeks* case] did not in fact make it more likely *ex ante* that Weeks would die, given the particular circumstances of the fire. For we know, although after the fact, that even if the hotel owner had been careful, Weeks would have died anyway.

WILLIAM M. LANDES & RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF Torts LAW* 235 (1987). As they implicitly acknowledge by referring to the owner’s “carelessness,” the owner was negligent given the *ex ante* risks. Their contrary conclusion is, again, based on *ex post* factual causation analysis masquerading as *ex ante* risk analysis. See *id.* at 238.

When the efficiency theorists acknowledge the distinct natures of the negligent conduct and actual causation requirements for tort liability, they generally define negligence as a departure from the economically optimal (aggregate social welfare maximizing) level of care and a cause as a strongly necessary (sine qua non or “but for”) condition. Given these definitions, assuming perfect information, risk neutrality, solvent defendants, and perfect enforcement (i.e., all valid claims are successfully pursued), and setting aside dynamic efficiency concerns (the effect on stability of entitlements of constantly shifting resources to their highest-value use, as contemplated by the efficiency theories of liability), a liability rule will be efficient if and only if it subjects every person who contributed to an injury liable for at least the full amount of the injury if she behaved inefficiently, relieves all but one person of any liability if they behaved efficiently, and holds (or leaves) that one remaining person (defendant or plaintiff) liable for the exact amount of the injury. Given the highly restrictive (implausible) assumptions, there are many theoretically efficient negligence liability rules, including those with no defense or a complete or partial (comparative responsibility) defense of plaintiff’s contributory negligence, with various options, including joint and several liability and proportional several liability, for allocating liability among multiple responsible parties. Strict liability rules will be efficient only if there is a contributory negligence defense and only if, ex ante, there is only one possible defendant. No liability rule will be efficient if any of the unrealistic assumptions are not satisfied.154

Moreover, given imperfect information, none of the rules proposed by the efficiency theorists are descriptively plausible.155 As most efficiency theorists admit,156 the courts cannot and do not define negligence as a failure to exercise economically optimal care. To employ such a standard, a court would need information on the costs and benefits of the actually taken precautions and all possible untaken precautions, by not only the potential defendant but also potential plaintiffs and anyone else in a position to


154. Brown, supra note 133; Wright, Allocating Liability, supra note 136, at 1170–77; Wright, New Old Efficiency Theories, supra note †, at 90–95.

155. See Wright, New Old Efficiency Theories, supra note †, at 83–90.

156. E.g., RESTATEMENT (THIRD) OF THE LAW OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 3 cmt. h (AM. LAW INST. 2010); LANDES & POSNER, supra note 152, at 20–21 (“Rarely will there be enough information about the costs and benefits of alternative safety measures to enable a confident judgment that the court’s solution is the efficient one.”); id. at 24 (same); Mark F. Grady, Proximate Cause and the Law of Negligence, 69 IOWA L. REV. 363, 385–91, 396–40, 396 n.68, 397 n. 69, 402–03 (1984); Hylton, Causation in Tort Law: A Reconsideration, supra note 153, at 104–07.
affect the ultimate outcome, and it would have to calculate the social costs of all the various possible combinations in order to find the optimal combination. Courts are never provided with such complete information, nor do they attempt to make such calculations.

A fundamental problem with the efficiency theories is that they assume that the courts should and do attempt to define liability rules to maximize aggregate social welfare, rather than to implement interactive justice, which focuses on the promotion of everyone’s equal external freedom in their interactions with others. As one of us has explained in detail elsewhere, the courts’ liability determinations are consistent with the justice theory rather than the efficiency theories. For example, regardless of the aggregate cost-benefit balance, the courts hold a defendant liable for putting others at risk if those risks were significant unless the risks were necessary for those being put at risk to obtain, directly as participants or indirectly as members of society, desired benefits from the defendant’s risk-creating activity that significantly outweigh the risks. Conversely, courts will not hold a defendant liable for failing to save another whom she did not put at risk unless, perhaps, the rescue effort would have imposed a minimal burden on the defendant, and will not hold a plaintiff contributorily negligent for putting herself at risk to save another unless there was no fair chance of saving the other.157

VII. CONCLUSION

As we previously noted, judges and lawyers frequently assert that philosophical analyses of causation are irrelevant and unhelpful in the law and, indeed, confuse rather than enlighten.158 While this is a fair description of much philosophical analysis, we have attempted in this paper to demonstrate that careful philosophical analysis is essential for distinguishing, clarifying and properly resolving the descriptive issue of causation (in its basic natural/actual/factual sense) and the normative issue of proper legal responsibility, each of which currently is poorly understood and thus the subject of confused and confusing discussions by courts and legal scholars.

It is true, as critics often state, that a proper philosophical analysis of causation leads to an almost infinite number of contributing conditions for any specific injury. However, while this is a problem for scientists, it should not worry lawyers. In law as well as ordinary life, we are not inter-

157. See Wright, Hand Formula, supra note 132, at 180–238; supra note 132 and accompanying text.
158. See supra note 3 and accompanying text.
ested in determining all of the practically innumerable immediate and remote causes of every event or state of affairs. Rather, we are only interested in the possible causes of discrete legally recognized injuries, and we are interested in only a very few of those possible causes: (i) the wrongful aspect of the defendant’s conduct, which is the proper focus of the analysis of causation in the law, (ii) the negligent aspect (if any) of the plaintiff’s conduct, which may result in application of the defense of contributory negligence, and (iii) highly unexpected and strongly necessary intervening conditions or independently strongly sufficient non-liable conditions, which generally will result in denials of legal responsibility despite tortious causation of the plaintiff’s injury by the defendant. Legal responsibility may also be denied for wrongful conditions that made only a trivial contribution to the injury or for an injury that was not the result of the realization and working out of a risk that made the defendant’s conduct wrongful.159 Moreover, failures to act to prevent an injury generally are not considered to be legally wrongful, or are considered wrongful only in very limited circumstances, if the person who failed to act had no pre-existing duty to act.160

While, pending further discussion, we disagree on some issues, such disagreement should not detract from our agreement on the most fundamental issues, including the NESS (weak necessity/strong sufficiency) covering law account of causation, the rejection of counterfactual possible worlds analysis, the inclusion of omissions and other absences as causes, the importance of focusing the causal analysis on the properties of events and states of affairs rather than events and states of affairs as a whole, and, relatedly, the importance of focusing the causal analysis for purposes of legal responsibility on the causal connection between the wrongful aspects of the defendant’s conduct and the relevant legal injury. Whatever disagreements we may have now or in the future pale in comparison to the defects of any alternative analysis of causation, especially the strong necessity (sine qua non, “but for”) analysis as an exclusive and/or counterfactual analysis or any account that purports not to rely on causal laws.

159. See supra text accompanying notes 6, 56–62, 100.
160. See Dobbs, supra note 48, at 853.