Legal Causation and Responsibility for Causing Harm

I POSIT THAT IN THE LAW, scientific knowledge is used to fix responsibility, not to reveal truth. It does this via its role in establishing legal causation. Thus, the role of “cause” in the courtroom has a different character than it does in epidemiology1 or toxicology. As with many other things in the law, legal cause is not a straightforward concept.

Take, for example, the following from Dobbs2:

... long but confusing legal tradition has assigned the term ‘causation’ to two entirely different kinds of legal problems. The first issue is about causation in the sense that term is used in everyday speech. Did the defendant’s negligent conduct cause the plaintiff’s harm or not? The second issue is about the appropriate scope of the defendant’s legal responsibility for negligent conduct that has in fact caused harm. This scope of responsibility issue often turns on whether the kind of harm plaintiff suffered was the same kind the defendant risked by his negligence.

As modern law books note, legal cause is composed of at least two distinct components, of which only one is related to “causality” as scientists understand it.

I argue that the alleged separation of the social issue of “scope of responsibility” from the objective question of scientific causality is illusory. “Scientific causality” also plays an essential social role in the legal setting. (Am J Public Health. 2005;95:S35–S38. doi:10.2105/AJPH.2005.061820.)

EXCLUDING SCIENCE FROM THE COURTROOM

THE ORIGINS OF LEGAL CAUSE

Late 19th and early 20th century legal thought was deeply affected by the historical setting of the industrial revolution, the rise of the corporation and the Trusts, and the resulting change in property relations, ownership, and the status of the individual. These changes in turn brought social turmoil and unrest. The railroad strike of 1877 produced a wave of urban armory construction designed specifically to quell civil unrest, and the Pullman strike, the Haymarket riot, and the unrest following the Depression of 1893 were all widely regarded as harbingers of revolution.

For the social classes who made the rules, state-enforced...
redistribution of wealth as a legal principle was the rotten fish on the dinner table. If the court—that is, the state—were to force the transfer of property from one private party to another in a law suit, it could only be justified in doing so by using rules of law that were apolitical and objective. According to Horwitz:

The idea of vindication of individual rights was intimately connected with the notion of objective causation. Only if it was possible to say objectively that A caused B’s injury would courts be able to take money from A and give damages to B without being charged with redistribution. Without objectionable causation, a court might be free to choose among a variety of possible defendants in order to vindicate the plaintiff’s claim. If the question of which of several acts caused the plaintiff’s injury was open to judicial discretion, how could private law stay clear of the dangers of the political uses of law for purposes of redistribution.

According to Horwitz, there was to be no room for judicial discretion or the intrusion of “policy” (widely understood to mean “politics”) in fixing causation historically; similar ideas are commonly expressed about science in the courtroom today.

But how could causation be “objectively” determined? The motivating idea was that the legal rules must be based on objective principles independent of any socially determined ends. The aim was the clichéd “rule of law, not of men.” Law, at least insofar as it dealt with private transactions between private parties, must remain neutral and untainted by politics. But in trying to find a standard for objective causation, orthodox legal theorists confronted the uncomfortable circumstance that the theoreticians and practitioners of objective cause, that is, philosophers and scientists, were advancing a view of “objective” causation that conflicted with the needs of “objective” causation in law.

For legal theorists of the late 19th century, moral forces and the choices of a freely acting subject were the bedrock of objective causation. As Horwitz put it, moral forces and free acts of human agents were “intelligible and objective a priori categories.” Objective moral forces proved a weak reed to lean upon. The dawn of the new century saw a growing skepticism concerning the objectivity of moral categories and a consequent weakening of the orthodox objectivist position. A series of blatantly pro-business and counter-intuitive court decisions, ostensibly based on this orthodox reasoning, proved a tempting and vulnerable target for progressive reformers. One of the most famous of these cases was Ryan v New York Central Railroad Company, wherein a railroad set fire to a house near one of its work sheds, which fire in turn set another house on fire and so on, until there was a general conflagration. The New York Supreme Court ruled that the railroad was liable only for setting fire to the first house, each fire thereafter not being proximate to the railroad.

This reasoning flew in the face of intuition and common notions of fairness, and became the poster child for judges hiding policy decisions behind the mask of orthodox objective causation. Nor was it the only such decision. In the last third of the 19th century and during the first two decades of the 20th century, many more court decisions made under cover of objective causation revealed the arbitrariness of the orthodox criterion and its biased application. In response, reformers advanced their own version of what distinguished legitimate antecedents of liability, that is, the foreseeability doctrine. If the consequences of a defendant’s conduct were foreseeable, then the conduct was within the scope of responsibility. This now commonly accepted idea, however, was strongly resisted by orthodox theorists:

If a house is properly built, if it is properly watched, if a proper fire apparatus is in operation, it can be prevented, when a fire approaches from a neighboring detached house, from catching fire. . . . A consequence of a foreseeability test would be that the capitalist would be obliged to bear the burden, not merely of his own want of caution, but of the want of caution of all who should be concerned in whatever he should produce.

Adopting foreseeability would allow courts to argue that carelessness on the part of downstream actors could be ignored even as liability is “traced back until a capitalist is reached. . . . If this law be good, no man of means could safely build a steam engine, or even a house.” Foreseeability would hasten the demise of capitalism itself.

It was a losing battle. The last gasp of legal orthodoxy was led by Harvard Law professors and legal “formalists” Christopher Columbus Langdell and Joseph Beale. Prefiguring Dworkin, the formalists believed there was a single notion of scientifically determined cause that could underwrite a judge’s decision to prevent a case from going to a jury. Abandoning moral categories, the formalists appealed to physical images that borrowed the prestige of the sciences. Writing in 1920, Beale set forth the following objective rules for determining this scientific causation:

The study of proximate causation will prove to be a study of activity of force or of risk. A result is usually created by the impulsion of an active force upon a passive force; . . . nothing but an active force can bring about that change of conditions which we call a consequence. . . . The whole problem may therefore be stated thus: when is one responsible for the operation (a) of an active force which he has created; (b) of an active force which acts upon a passive force which he created, or upon a passive force which he was legally bound to change. Each time one or more active causes operate on a condition to create a new condition, a new causal step is taken, ending with the given result. This result is the direct result of the active force or forces which last acted upon the immediately precedent condition, and is the indirect result of the earlier acting forces. The active force which brings about the result without the intervention—the subsequent coming into action—of any other force is a direct cause of the result.

The following is how Beale’s mechanical metaphor described the limits of scope of responsibility: where the defendant’s active force has come to rest in a position of apparent safety, the court will follow it no longer; if some new force later combines with this condition to create harm, the result is remote from the defendant’s act.

That these “objective rules” were of little help was obvious when judges tried to put them into practice. By the 1930s, objective causation as previously understood was all but dead. Foreseeability had become the standard for scope of responsibility. In the process, scope of responsibility had become transformed from an alleged objective natural sequence to a policy question of how risks were to be distributed.

**OBJECTIVE** CAUSATION TODAY

The tactic of separating the old objective cause into two parts—the value-laden part of proximate...
cause and a supposedly value-free residual—produced the instant resurrection of the residue of cause-in-fact as the new objective form of scope of responsibility. The reason this maneuver seemed unproblematic at the time is related to the fact that the general causation questions formerly at issue in tort cases involved well-understood and familiar principles, so that the only “scientific” question was whether a specific case was an instance of a general well-understood principle or not. After that, the legal-cause question was concerned with the traditional problem of determining scope of responsibility as formulated from the 1920s onward.

Law books, as in various editions of Prosser’s standard treatise on the law of torts, typically express this idea in the following way: “Causation is a fact. It is a matter of what has in fact occurred. A cause is a necessary antecedent: in a very practical sense the term embraces all things which have so far contributed to the result that without them it would not have occurred.” The acceptance here of the view that all antecedents are relevant to cause-in-fact (a view that goes back to John Stuart Mill and which was rejected by orthodox legal writers because it appeared to leave the vacuum cleaner company liable) forced mainstream 20th century theorists to concentrate on proximate cause to the neglect of deeper analysis of cause-in-fact. The hard cases seemed to involve determination of the scope of responsibility. By contrast, the allegedly factual residue, that is, cause-in-fact, seemed commonsensical and easy. A notion of cause-and-effect seemed the very precondition of rational thought about the universe, every part of which “is effect and cause in the same sense in which those parts are with which we are most familiar.”

To understand the meaning of objective cause-in-fact in the legal process, we need to ask what function it plays in the tort system. The following three things need to be shown to prevail in a negligence suit:

• First, the existence of negligent conduct by the defendant.
• Second, the existence of damages by the plaintiff.
• Third, the showing that the defendant’s conduct was the legal cause (both proximate and cause-in-fact) of the plaintiff’s harm.

If the plaintiff can persuade the jury of these three elements, then he or she is entitled to compensation (usually monetary). Compensation in this context is a state-enforced redistribution of wealth. In our society, this is a drastic remedy. One might argue that it is among the most drastic the state can exact. I am not arguing that this scale of values is justified, much less rational, but only that it prevails at this time and place (hope springs eternal). What do we require to justify it?

Consider a worker exposed to the asbestos insulation of two companies. Both asbestos products were used interchangeably, neither had warning labels, and both companies were aware of the hazards of their product even as tens of thousands of workers used them over a span of decades. After many years of exposure, the worker is diagnosed with mesothelioma, a uniformly fatal and painful cancer of the lining of the lungs that is associated almost entirely with asbestos exposure. Because most cancers start with the initiation of malignancy in a single cell, it is almost certain that a particular fiber from one or the other of the companies began the process. But which one? We don’t know, because exposure occurred to both products in the relevant time period. Should both companies therefore be let off because we cannot prove that either of them actually “caused” the disease? This doesn’t seem right and, indeed, under current law, both companies can be held liable, even though only one actually “caused” the disease in a strict sense.

Now suppose, prior to the trial, one of the companies sponsored an animal study showing that its asbestos fiber type was not the cause of the mesothelioma because its shape made it unable to get to the periphery of the lung where the cancer arose. They admit their fiber is even more likely than their co-defendant’s to cause lung cancer, but this was not the kind of cancer the plaintiff had. If the jury members were persuaded by the scientific findings, then the defendant would be found not liable.

Now consider a third case. The victim is an inventor who is developing a new asbestos product. In the course of research and development, he is exposed to asbestos and develops mesothelioma years later. Now, there is no second party and no one to sue. The question of “causation” is irrelevant. But the same biological causation process occurred in all three cases. The worker’s lung had no idea of the differences in circumstances in the three examples.

With these examples in mind, we return to the question of justification of compensation in the tort system. First, we can easily dismiss the following two suggestions: (1) that compensation is a penalty exacted on those who are negligent, and (2) that compensation is to make whole someone who suffered an undeserved harm. The answer cannot be that compensation is due because of negligent conduct, because negligent conduct that doesn’t “cause” harm is not liable. If you drive inattentively and just miss hitting someone, you are not liable. Nor can the justification be because the victim suffered an undeserved harm, because many people suffer undeserved harm and are not compensated by anyone. But the conduct and the harm together might be the reason. Legal cause provides the connection. Indeed, it permits us to identify specific actors among all actors (e.g., specific chemical companies that have acted negligently among all chemical companies that have acted negligently) and specific victims of undeserved harm (among all victims of undeserved harm) and tie them together. When a link of causation is shown, we say one party is “at fault,” or responsible. Cause-in-fact in a tort case is another way of defining fault or responsibility. It is not a logical requirement, because it is neither necessary nor sufficient for the requirement of monetary exchange between the parties. Nor does it have moral relevance, which is something I shall argue in a moment. It only has meaning in the legal context with respect to some social sense of what it means to be responsible for a harm. Importantly, it is quite a different concept than Rothman’s notion of causation, which does not carry this additional freight.

Consider our examples. In the first example, neither company could be shown to be the “cause” of the cancer, although one of them surely caused it. Yet both were considered liable for purposes of compensation. In the second example, only the one company was considered “at
fault," even though both companies behaved equally badly, and it was only "good luck" for one company that the other company’s product killed the worker before its product had a chance to do it. If "fault" in this sense has a moral relevance, it is hard to see the moral difference between the two companies, which are distinguished from one another only by luck. In the third example, fault is not an issue and so neither is cause.

Cause-in-fact seems neither necessary, nor sufficient, nor morally relevant. Indeed, cause-in-fact in the scientific meaning is only relevant in any sense in the legal setting when it is useful for establishing responsibility. "Responsibility" is another word for who must assume the burden. It is a social decision about how costs will be distributed. Thus, the "scope of responsibility" question, though banished by the distinction of proximate cause, is still present in another form in cause-in-fact. The social continues to lurk in objective causation, where it is now a sign of fault or responsibility, with all the connotations those words carry.

About the Author
The author is with the Boston University School of Public Health, Boston, Mass.

Requests for reprints should be sent to David Ozonoff, MD, MPH, Boston University School of Public Health, 715 Albany Street, Boston, MA 02118 (e-mail: dozonoff@bu.edu).

This commentary was accepted on January 27, 2005.

Acknowledgments
This work was supported in part by the Project on Scientific Knowledge and Public Policy.

The author thanks D. Michaels and L. Boden for their helpful comments.

References