

Fordham Law School

FLASH: The Fordham Law Archive of Scholarship and History

Faculty Scholarship

2002

Unrealized Torts

Benjamin C. Zipursky

Fordham University School of Law, bzipursky@law.fordham.edu

John C.P. Goldberg

Harvard Law School, jgoldberg@law.harvard.edu

Follow this and additional works at: https://ir.lawnet.fordham.edu/faculty_scholarship



Part of the [Law Commons](#)

Recommended Citation

Benjamin C. Zipursky and John C.P. Goldberg, *Unrealized Torts*, 88 Va. L. Rev. 1625 (2002)

Available at: https://ir.lawnet.fordham.edu/faculty_scholarship/834

This Article is brought to you for free and open access by FLASH: The Fordham Law Archive of Scholarship and History. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of FLASH: The Fordham Law Archive of Scholarship and History. For more information, please contact tmelnick@law.fordham.edu.

VIRGINIA LAW REVIEW

VOLUME 88

DECEMBER 2002

NUMBER 8

ARTICLES

UNREALIZED TORTS

John C.P. Goldberg & Benjamin C. Zipursky***

INTRODUCTION.....	1626
I. REALIZED WRONGS	1636
A. <i>Crime versus Tort</i>	1636
B. <i>Tort as Civil Recourse</i>	1641
II. WHEN IS HEIGHTENED RISK A COGNIZABLE INJURY?	1650
III. RISK OF FUTURE INJURY AND THE LAW OF EMOTIONAL DISTRESS.....	1660
A. <i>The Current State of Emotional Distress Doctrine</i>	1660
1. <i>Basic Doctrine</i>	1660
2. <i>Fear-of-Disease Doctrine</i>	1664
B. <i>The Current Conceptualization of Emotional Distress Claims</i>	1667
C. <i>Against NIED</i>	1672

* Professor, Vanderbilt University Law School.

** Professor and Associate Dean for Academic Affairs, Fordham University School of Law.

The Authors thank Professor Kenneth Abraham and the participants in the conference on Liability for Inchoate and Future Losses for very helpful discussions of an earlier draft of this paper, particularly Professor Geistfeld for his extended commentary. This Article also benefited from discussion at faculty workshops held at Fordham University School of Law and the University of Illinois College of Law. Mark Brandon, Lisa Bressman, Allison Danner, Steve Hetcher, Gail Hollister, Nancy King, Richard Nagareda, Gideon Parchomovsky, Robert Rasmussen, Tony Sebok, Catherine Sharkey, Suzanna Sherry, Peter Siegelman, Steve Thiel, and Mike Vandenberg offered many helpful comments, and Neera Chatterjee and Danielle DePalma provided excellent research assistance. Our work was supported by generous research assistance from Vanderbilt University Law School and Fordham University School of Law.

D. Rethinking NIED: Agency and Responsibility.....	1676
1. Fraud and Subjectivity.....	1678
2. Revisiting Exceptions to the Default Rule of No Duty.....	1685
3. Evading the Dilemma of Fear-of-Future-Disease Claims: Secondary Legal Qualities.....	1688
E. Reconceiving Fear-of-Future-Disease Claims.....	1693
IV. MEDICAL MONITORING CLAIMS	1701
A. Uncertainty over the Nature of Medical Monitoring Claims	1702
B. Medical Monitoring and the Duty to Aid the Imperiled ..	1709
CONCLUSION.....	1715

INTRODUCTION

CURRENTLY, lawyers and courts are struggling with tort suits brought by plaintiffs who seek compensation even though they have not suffered any bodily harm or manifested symptoms signaling the onset of a disease. These include, most prominently, claims that a defendant's carelessness exposed the plaintiff to known carcinogens such as asbestos, or known pathogens such as HIV, thereby putting the plaintiff at risk of developing a serious or fatal illness in the future. As public awareness and concern over environmental toxins and other hazardous substances continue to grow, there is every reason to suppose that these claims will continue to be pressed in courts around the country.¹ Yet judges and lawyers so far have proceeded without a firm grasp of how this area of the law is unfolding or ought to unfold.

By asking conferees to address this active and unsettled area of tort law, the conference giving rise to this volume poses a basic question to tort scholars: How can we best contribute to the analysis of an emerging area of tort law? Perhaps the standard response to this question invokes a sentiment sometimes attributed to Justice Holmes. The life of the law, he famously proclaimed, has not

¹ The prospect of a continuing flood of claims for fear of cancer by those exposed to asbestos has moved the Supreme Court to review an unpublished state court decision permitting such a claim. See *Norfolk & W. Ry. Co. v. Ayers*, 122 S.Ct. 1434 (2002) (mem.) (granting petition for review). We discuss *Ayers* briefly below. See *infra* notes 128, 182.

been “logic” but “experience.”² It follows, according to his modern counterparts, that legal scholars “add value” by getting beyond the concepts of legal doctrine to the real or root policy question of whether, or under what circumstances, recognition of these claims will achieve desired social ends.³ Thus, what is called for on the Holmesian view is not the sort of formal or conceptual analysis (“logic”) that occupies lawyers and judges, but pragmatic analysis (whether microeconomic or more broadly welfarist) of what ends might be served by recognizing or not recognizing these claims.

There is, however, a different approach to scholarly analysis in unsettled areas of law, one traceable to another titan of American law: Justice Cardozo.⁴ The Cardozoan approach rejects the Holmesian dichotomy between logic and experience. Indeed, it maintains that scholars—at least scholars writing for lawyers and judges—do more harm than good by attempting to answer questions of the sort posed by “future injury” cases without attending to the conceptual issues that they raise. Cardozoan “pragmatic conceptualism,” as we have previously called it, shares the Holmesian goal of promoting ongoing revision of law and legal practices to meet new challenges, like those posed by toxic exposure litigation.⁵ Yet it regards the concepts of legal doctrine as an aid, rather than a hindrance, to that progressive project.

In past works, we have made a case for pragmatic conceptualism over pragmatic instrumentalism as a means of understanding tort

² Oliver Wendell Holmes, Jr., *The Common Law* 1 (Boston, Little, Brown, & Co. 1948) (1881).

³ See John C.P. Goldberg, *Twentieth Century Tort Theory*, 90 *Geo. L.J.* (forthcoming 2003) (manuscript on file with the Virginia Law Review Association) (describing Holmesian compensation-deterrence theory in these terms).

⁴ See, e.g., John C.P. Goldberg, *The Life of the Law*, 51 *Stan. L. Rev.* 1419, 1461–73 (1999) (arguing that Justice Cardozo employed a pragmatic form of conceptualism in his decisionmaking); John C.P. Goldberg & Benjamin C. Zipursky, *The Moral of MacPherson*, 146 *U. Pa. L. Rev.* 1733, 1812–24 (1998) (articulating Justice Cardozo’s belief that one could deploy the concept of duty in negligence law in a manner that was progressive and pragmatic without being an instrumentalist); Benjamin C. Zipursky, *Pragmatic Conceptualism*, 6 *Legal Theory* 457, 474–78 (2000) (outlining the jurisprudential position of pragmatic conceptualism).

⁵ See Zipursky, *supra* note 4; see also Goldberg, *supra* note 4, at 1462.

law.⁶ This puts us distinctly in the minority: The Holmesian conception of legal scholarship has dominated among torts scholars in the last century. Yet, one occasionally finds precedents that attest to the efficacy of the minority approach. An unlikely example is provided by the famous article entitled "Privacy," published in 1960 by Dean Prosser, a confirmed Holmesian.⁷ "Privacy" aimed to rationalize seemingly disparate judicial decisions permitting liability to attach to invasions of the "right 'to be let alone'"—a right that Warren and Brandeis had themselves earlier extracted from the case law.⁸ Judicial issuance of these decisions, Prosser commented, "has gone on without any plan . . . [and] without much realization of what is happening or its significance."⁹ Still, he maintained, one could discern order amidst the confusion: Privacy claims could be reduced to one of four archetypal forms.¹⁰ By prompting this realization, Prosser hoped to clarify lawyerly thinking about the nature and proper scope of liability that attaches to claims for tortious invasion of privacy. In fact, his re-conceptualization of the cases—which he wrote into the Second Restatement—has set the terms of legal analysis in this area ever since.¹¹

Our goal in this Article, although more ambitious than doctrinal brush clearing, is not dissimilar to that of Prosser in "Privacy." We aim to aid the courts in realizing what is happening in the emerging

⁶ In addition to the works cited in note 4, see John C.P. Goldberg & Benjamin C. Zipursky, Concern for Cause: A Comment on the Twerski-Sebok Plan for Administering Negligent Marketing Claims Against Gun Manufacturers, 32 Conn. L. Rev. 1411, 1418–23 (2000) [hereinafter Concern for Cause]; John C. P. Goldberg & Benjamin C. Zipursky, *The Restatement (Third) and the Place of Duty in Negligence Law*, 54 Vand. L. Rev. 657, 723–36 (2001) [hereinafter *The Place of Duty*]; Arthur Ripstein & Benjamin C. Zipursky, Corrective Justice in an Age of Mass Torts, in *Philosophy and the Law of Torts* 214, 231–44 (Gerald J. Postema ed., 2001); Benjamin C. Zipursky, Legal Malpractice and the Structure of Negligence Law, 67 Fordham L. Rev. 649, 675–76 (1998).

⁷ William L. Prosser, Privacy, 48 Cal. L. Rev. 383 (1960).

⁸ Samuel D. Warren & Louis D. Brandeis, The Right to Privacy, 4 Harv. L. Rev. 193, 195 (1890).

⁹ Prosser, *supra* note 7, at 422.

¹⁰ *Id.* at 389 (describing four distinct privacy invasions as: intrusion upon plaintiff's privacy, "public disclosure of embarrassing private facts," publicly placing plaintiff in a false light, and the use of the plaintiff's "name or likeness").

¹¹ See Restatement (Second) of Torts § 652A (1977); Sheldon W. Halpern, *The Law of Defamation, Privacy, Publicity and Moral Right: Cases and Materials on Protection of Personality Interests* 412 (4th ed. 2000) (noting that courts frequently employ Prosser's categories).

and important class of tort cases now gathered under the loose banners of "future injury" or "inchoate torts." More specifically, we aim to illuminate and help guide tort practice by attending to a set of theoretical and conceptual questions raised by these cases.

The first and most theoretical question they pose concerns the role of "injury" in tort law. To the extent that courts allow recovery in suits based on not-yet-realized injuries, they might suggest the demise of the traditional requirement that a tort plaintiff—or at least a negligence or products liability plaintiff—prove that she suffered an injury at the hands of a wrongdoer. On this conception, as the title of this conference suggests, the imposition of liability in heightened-risk cases is akin to the imposition of punishment for inchoate crimes such as attempts and conspiracies. Thus, in terms of theoretical significance, these cases might be taken to signal a substantial break from traditional distinctions between crime and tort as well as between public and private law.

Although theoretical, the issue of the injury requirement in tort is quite important at a pragmatic level. Still, we concede that it is likely not the most pressing issue in the minds of lawyers and judges facing claims for heightened risk of physical injury. Rather, they are more likely to be wrestling with the issue of whether or when the imposition of such risk constitutes a compensable injury to the person(s) exposed to the risk. On this view, the issue is not whether there are any "inchoate" torts, but when a risk can be said to have ripened into a cognizable harm.

This question, unfortunately, is hardly less daunting because it leaves jurists with the difficult task of defining the injury suffered by plaintiffs who have been exposed to toxic substances yet have not manifested a physical injury or illness. Indeed, one tends to encounter two very different accounts on this score. The first maintains that a heightened risk of bodily harm or illness is a cognizable injury in and of itself. That is, regardless of whether the risk has matured into bodily, emotional, or economic harm, or ever will, exposure to an increased risk of those harms is an injury in principle sufficient to support a claim for negligence or some other tort. Therefore, a plaintiff who can establish that another actor negligently exposed her to a known carcinogen or pathogen has made out the elements of the *prima facie* case: duty, breach, cause, and injury. Although, on this conception, heightened-risk claims are *not*

claims for inchoate torts, they may appear quite close to being such claims by expanding the definition of injury to include heightened risk.

The alternative account of heightened risk as injury maintains that exposure to heightened risk of bodily harm or disease is a genuine injury, but only when it is linked in the right way to some other harm experienced by the plaintiff. The linked harms most commonly identified by this sort of argument are emotional distress and economic loss. A heightened risk of future injury often is tied to a present fear or apprehension in the plaintiff that arises out of exposure to the risk. On this view, future injury cases in which recovery is permitted are not claims for risk exposure, but instead claims that fall within, or at least in the vicinity of, the category of claims for negligent infliction of emotional distress. A heightened risk of future injury also may be connected to present economic loss: Those wrongfully exposed to toxic substances may reasonably incur expenses associated with monitoring their physical condition in light of their exposure. Thus, some plaintiffs treat heightened-risk claims as if they were cousins of claims made against an accounting firm by investors who have suffered economic loss because of the firm's negligently prepared audit report.

Analyses that attempt to hitch heightened-risk claims to the injuries of emotional distress and/or economic loss are more grounded in familiar doctrine than those claiming that heightened-risk claims are for inchoate wrongs, or that heightened risk is injury by itself. They are hardly unproblematic, however, as can be illustrated by brief consideration of a leading case in this area, the United States Supreme Court's decision in *Metro-North Commuter Railroad Co. v. Buckley*.¹²

Buckley was exposed to substantial amounts of asbestos dust in his job as a pipefitter for Metro-North.¹³ Although symptomless, he sued the railroad on the ground that its negligence (for example, in failing to provide safety equipment) had caused him to experience a one to five percent increase in the likelihood of contracting cancer and/or asbestosis.¹⁴ The suit was brought under the Federal

¹² 521 U.S. 424 (1997).

¹³ *Id.* at 427.

¹⁴ *Id.*

Employers' Liability Act ("FELA"), which permits a railroad worker to recover for an injury resulting from the negligence of his employer.¹⁵ Buckley's lawyers apparently did not argue that a FELA plaintiff may recover without proof of injury. Nor did they argue that exposure to heightened risk of disease is, in and of itself, an injury. Rather, they framed Buckley's claims under FELA as, respectively, sounding in negligent infliction of emotional distress and negligence causing economic loss.

Buckley's lawyers had some grounds for optimism regarding the emotional distress claim. The Supreme Court had held three years earlier in *Consolidated Rail Corp. v. Gottshall* that a railroad worker could recover for emotional distress caused by her employer's carelessness if she "*sustain[ed] a physical impact as a result of a defendant's negligent conduct, or . . . [was] placed in immediate risk of physical harm by that conduct.*"¹⁶ Although Buckley had not been placed in immediate risk of physical harm by Metro-North's negligence, his lawyers argued that he was nonetheless entitled to recover under the first prong of *Gottshall* because his distress over his increased risk of disease was attendant on the "physical impact" of asbestos fibers lodging in his lungs.¹⁷

Writing for a majority of seven, Justice Breyer rejected this claim. The "physical impact" language of *Gottshall*, he reasoned, was not meant to establish a separate prong under which FELA plaintiffs could establish a claim for negligent infliction of emotional distress.¹⁸ Rather, it merely referred to the well-established—if not obviously justified—doctrine holding that a plaintiff who suffers a traumatic physical harm (for example, a broken limb, a laceration, or an illness) may recover emotional distress damages as parasitic on the predicate physical injury.¹⁹ FELA plaintiffs bringing pure emotional distress claims—distress not predicated on a physical injury—must establish, as Buckley could not, that the defendant's negligence had placed them in the "zone of danger," that is, at imminent risk of physical harm. In offering this arguably nar-

¹⁵ 45 U.S.C. § 51 (2000).

¹⁶ 512 U.S. 532, 547–48 (1994) (emphasis added).

¹⁷ *Buckley*, 521 U.S. at 427–28.

¹⁸ *Id.* at 431–32.

¹⁹ *Id.* at 429–30; see also Dan B. Dobbs, *The Law of Torts* § 302, at 822 (2000) (stating the traditional rule).

row construction of the statute, the Court relied heavily on "policy reasons" immediately familiar to students of the common law of torts.²⁰ To permit claims such as Buckley's, the Court reasoned, would be to expose FELA defendants to "unlimited and unpredictable liability," to present the courts with hard-to-adjudicate emotional distress claims, and to invite invalid or "trivial" claims.²¹

Buckley's lawyers argued in the alternative that he would have to undergo, as a matter of sound preventive medicine, regular medical tests to monitor the development, if any, of the diseases at which he was now at a higher risk of contracting.²² Because Buckley would not have incurred these medical monitoring expenses but for Metro-North's negligence, his attorneys argued that the company's negligence had caused Buckley a form of economic loss.²³ Accordingly, Buckley's counsel sought damages equal to the expected future cost of a reasonable program of medical monitoring.²⁴

In contrast to its emphatic rejection of Buckley's emotional distress claim, the Supreme Court displayed ambivalence in addressing his medical monitoring-economic loss claim. The Court conceded, as it had not earlier in the opinion, that Buckley "*has* suffered wrong at the hands of a negligent employer."²⁵ Still, the Court was troubled by the "potential systemic effects" of recognizing a claim such as Buckley's.²⁶ Such recognition, the majority speculated, might invite "tens of millions" of claims, many of which might be fraudulent, for relatively minor injuries.²⁷ Especially in the context of asbestos litigation, where the assets available to compensate tort victims are not sufficient to permit the recovery of full compensation by all those injured, the Court worried that permitting Buckley's claim would set inequitable priorities among claimants, permitting the exposed-but-uninjured to recover before (and possibly more than) those who have manifested diseases.²⁸

²⁰ *Buckley*, 521 U.S. at 433.

²¹ *Id.* (quoting *Gottshall*, 512 U.S. at 557).

²² *Id.* at 438.

²³ *Id.*

²⁴ *Buckley v. Metro-North Commuter R.R.*, 79 F.3d 1337, 1346-47 (2d Cir. 1996), *rev'd*, 521 U.S. 424 (1997).

²⁵ *Buckley*, 521 U.S. at 443.

²⁶ *Id.*

²⁷ *Id.* at 442-44.

²⁸ *Id.* at 435-36.

In light of these concerns and its prior holding on emotional distress, one might have expected the majority to hold that claims for medical monitoring expenses are simply not cognizable under FELA. Yet, Justice Breyer's opinion stopped short of that blanket rule. Courts, he concluded, cannot recognize suits seeking lump-sum payments for future medical monitoring costs. Nonetheless, judges may be able to order the creation of funds that reimburse plaintiffs' reasonable monitoring expenses on an ongoing basis. Thus, the majority remanded the case for the Second Circuit Court of Appeals to consider whether FELA "might, or might not, accommodate medical cost recovery rules more finely tailored than the rule we have considered."²⁹

Justice Ginsburg, in dissent, described this final twist in the majority's analysis as "enigmatic."³⁰ The policy concerns identified by the majority, she reasoned, either do or do not counsel against recognizing Buckley's cause of action for pure economic loss.³¹ If they do, then he has no right to any recovery. If they do not, then he has such a right, and the form of the remedy appropriate to vindicate that right is of no particular legal significance.

Buckley indicates that heightened-risk litigants (and courts adjudicating their claims) do not face a smooth road simply by steering clear of the doctrinal and theoretical potholes attending efforts to argue that exposure to risk is an injury in and of itself. This should not be a surprise. Even outside the context of heightened-risk claims, the law concerning negligent infliction of emotional distress and pure economic loss is confused. In the former category, for example, courts employ, or have employed, a tangled array of concepts such as "predicate injury," "parasitic damage," "impact," "zone of danger," "foreseeability," "direct/indirect," and "bystander."

In short, we see in modern tort litigation over conduct causing heightened risk of future injury what Dean Prosser saw in the early privacy cases: lawyers and judges muddling through, without a full realization of what is happening or its significance. The goal of this Article, like Prosser's goal in "Privacy," is to help them move toward a clearer realization of what is happening in this important,

²⁹ Id. at 444.

³⁰ Id. at 445, 455 (Ginsburg, J., dissenting).

³¹ See id. at 448.

emerging area of tort law. We will proceed in that project by analyzing the issues posed by future injury cases in the order just discussed.

Parts I and II will address the theoretical question of the significance of future injury claims for tort law as a whole.³² Part I will maintain that future injury claims should not be regarded as pointing toward the recognition of tort liability for inchoate wrongs. To the contrary, a tort plaintiff must establish the occurrence of a realized wrong before liability may attach. This realization requirement is central to tort law and distinguishes it as a species of private law from the public law departments of criminal and regulatory law.

Part II will entertain the argument that heightened risk of physical injury, in and of itself, is a cognizable harm. Here, we will assume that exposure to heightened risk of physical injury can constitute a loss of welfare. That is, we will take seriously the idea that claims for heightened risk are claims for "choate" rather than "inchoate" torts, with the harm being the heightened risk of injury itself. Nonetheless, we will argue that this sort of harm usually does not, as a matter of existing negligence law, suffice to generate a tort cause of action in the plaintiff. In this sense, it is, at most, a non-cognizable harm—a harm but not an injury. This aspect of the law is grounded in the particular conception of duty that undergirds negligence. The duties typically recognized within the law of negligence are duties to take care not to cause "ultimate" or "target" harms, such as bodily harm or illness. By contrast, negligence law does not treat the "intermediate" or "unripened" harm of heightened risk as actionable injury. We will further suggest that tort law has sound reasons for generally recognizing duties to avoid causing ultimate harms, as opposed to duties to avoid causing the intermediate harm of enhanced risk of physical harm.

Part III will turn from tort theory writ large to the more immediate doctrinal issue of whether future injury suits can sometimes properly be cast as claims for negligent infliction of emotional distress. Here, we will argue that conceptualizing heightened-risk

³² Although this analysis is integral to the analysis that follows in subsequent Parts, those readers most interested in the practical-doctrinal issues raised when future injury cases are framed as claims for emotional distress or economic loss may wish to proceed to Parts III and IV.

claims in terms of emotional distress, although understandable, is a mistake, in part because the whole category of negligent infliction of emotional distress is infested with confusions. After offering a fresh interpretation of this much-analyzed-yet-misunderstood area of negligence law, we will argue that insofar as they are actionable because they are attended by an injury having to do with the victim's apprehension of the risk, heightened-risk claims are better conceived of as close cousins of torts such as assault, nuisance, and false imprisonment. That is, the ultimate injury attending these heightened-risk claims—and indeed most claims falling under the ill-conceived rubric of negligent infliction of emotional distress—is a harm to the plaintiff's interest in being able to function in society free of the pall of a significant threat to his well-being created by the wrongful acts of others. Conceiving of the injury in these cases as being placed under threat, we will argue, accomplishes what Prosser accomplished when he parsed privacy claims into four distinct privacy torts: it better captures the interests at stake, thereby permitting courts to draw more intelligent lines in determining when liability should attach in negligence for such injuries.

Part IV will argue that the “enigmatic” last step in *Buckley*'s analysis of claims for heightened risk—as claims for the expenses associated with medical monitoring—correctly linked the issues of right and remedy, although for reasons the Court itself seems only dimly to have perceived. The majority was right to conclude that the form of remedy—lump-sum versus scheduled payments—is critical to the viability of claims like *Buckley*'s, although not for the reasons it offered. Rather, the form of payment is important because these claims are not rightly conceptualized as seeking compensation for a completed tort. Instead, they are requests for courts to exercise their equitable powers to enjoin parties like Metro-North to undertake affirmative efforts to protect plaintiffs whom have been exposed to heightened risk of disease by their negligence. In other words, in our view, *Buckley*'s second claim did not seek a compensatory remedy for a realized wrong; rather, it sought a proactive court order requiring Metro-North, whose negligent conduct imperiled *Buckley*, to take affirmative steps to “rescue” *Buckley* (minimize his peril) by paying for his medical monitoring.

We will conclude by returning to the methodological and jurisprudential issues with which we commenced this introductory discussion. As the foregoing roadmap indicates, our approach to the problem of liability for future injuries stands in contrast to the Holmesian-instrumentalist approach employed by the majority of leading torts scholars, including many of the distinguished panelists at this conference. By applying pragmatic conceptualism to a contemporary problem in torts, we hope to supplement previously presented philosophical, historical, and doctrinal arguments for this approach with a tangible demonstration of its worth.

I. REALIZED WRONGS

A. *Crime versus Tort*

There is a fundamental distinction between criminal law, on the one hand, and tort law, on the other.³³ Criminal law sometimes prohibits and punishes genuinely inchoate wrongs—uncompleted wrongful acts. Tort law does not.

Suppose *X* undertakes a plot to avenge a perceived slight by *Y* by blowing up *Y*'s car in the middle of the night when no one is in it. *X* carefully records his plan in his diary, constructs the explosive device, then drives his car (with diary, device, and a map marking the route on the front seat) toward *Y*'s house. However, *X*'s plot is foiled when he is pulled over by the police one block from *Y*'s house for a traffic violation, at which point his plot is discovered.

In the majority of jurisdictions, the fact that *X* has failed in his effort does not defeat criminal liability,³⁴ nor does the fact that the once-significant threat to *Y*'s property has now passed. So long as *X* can be found to have acted with the requisite intent, and to have made a "substantial step" toward completion of the intended crime, failure to complete the plan does not decriminalize the conduct; it instead alters the charges that the prosecution may bring from, say, arson or criminal destruction of property to attempted

³³ Cf. *Palsgraf v. Long Island R. Co.*, 162 N.E. 99, 101 (N.Y. 1928) (explaining that to permit a tort plaintiff to sue for acts that are wrongful only to others is to "ignore the fundamental difference between tort and crime").

³⁴ Wayne R. LaFare, *Criminal Law* § 6.2, at 550 (3d ed. 2000).

arson and attempted destruction of property.³⁵ Likewise, if in the foregoing hypothetical Z helped plot the bombing, X and Z may, under the right circumstances, be prosecuted for conspiracy to commit arson, and so on.³⁶

Despite the fact that X and Z have committed crimes, Y cannot, on these facts, sue either of them in tort. Claims for battery, assault, and conversion will fail because they require Y to have suffered, respectively, harmful or offensive physical contact, the apprehension that such contact was imminent, or actual damage to property. There are no torts of attempted battery, attempted assault, or attempted conversion. Similarly, after X's and Z's plot has been foiled, Y has no cause of action against either for conspiracy because, strictly speaking, there is no tort of civil conspiracy; rather, that phrase refers to a rule of joint liability for co-tortfeasors.³⁷

The same holds true for garden-variety claims sounding in negligence and products liability. Take a classic negligence case such as *Vaughn v. Menlove*.³⁸ Menlove built his haystack in a maimer that increased the risk that it would catch fire, and hence increased the risk that Vaughn's neighboring property and dwellings would be damaged by fire.³⁹ Suppose, however, the hay had never caught fire and was later disintegrated. At that point, Vaughn would not have been able to sue Menlove in negligence for building a haystack that at one time increased Menlove's risk of suffering property damage.⁴⁰

³⁵ Model Penal Code § 5.01(1), at 295-96 (1985) (defining attempt). Alternatively, X might be charged with a possession crime. LaFave, *supra* note 34, § 6.2, at 537-38.

³⁶ Model Penal Code § 5.03(1), at 382-83 (defining conspiracy).

³⁷ Dobbs, *supra* note 19, § 340, at 936 ("Conspiracy is not a tort in itself; it reflects the conclusion that each participant should be liable for the tortious course of conduct.").

³⁸ 132 Eng. Rep. 490 (C.P. 1837).

³⁹ *Id.* at 491.

⁴⁰ To constitute a nuisance, an activity ordinarily must "interfere with" another's use of her property by, for example, generating noxious odors or noise. However, some courts have held that a noninvasive hazard on a property can constitute a nuisance to neighboring property owners, albeit only if it poses such significant dangers as to render the neighbors' property essentially useless and valueless, as might the construction and operation of a fireworks or explosive factory next to a residence. See 1 Dan B. Dobbs, *Law of Remedies: Damages-Equity-Restitution* § 5.6(1), at 754 (2d ed. 1993); Dobbs, *supra* note 19, § 466, at 1332. Menlove would have had no such claim against Vaughn.

Likewise, suppose car driver *D* speeds down a residential street for amusement, indifferent to the well-being of others, passing *P*'s yard in the process. *P* has her back turned to the street and is listening to a portable CD player and thus does not notice *D* driving by. Later, however, she sees police cars driving by her yard. Upon investigation, *P* surmises that *D* in fact drove his car at a very high rate of speed in close proximity to her while she was working in the yard. *P* has no cause of action in negligence simply because *D*'s negligent driving at one time increased her risk of injury, even though the act may perhaps have constituted the crime of reckless endangerment.⁴¹ Or suppose that mall owner *M* fails for a period of six months to take reasonable precautions to secure the mall's parking lot from criminal activity. After that time, the mall closes. Even in a jurisdiction recognizing an obligation owed by *M* to provide reasonable security for the benefit of mall patrons, those patrons cannot prevail in a negligence suit by establishing by convincing evidence that *M*'s carelessness increased the risk to each of them of being attacked.

Each of the foregoing examples involves conduct that constitutes a departure from legal standards of required conduct. Moreover, in each example, a risk of injury was posed to another person. Yet in each situation the risk was not realized with respect to that person. These examples indicate that, although tort law concerns itself with wrongs, not every wrong or even every legal wrong is a tort. Tort utilizes a conception of wrongs that involves "realized" wrongs. If a physician negligently fails to ascertain whether her patient is Rh- before delivering her baby, and the patient is in fact Rh+, yet that mistake has no significance for the care accorded to mother and child, no tort has occurred. The physician's conduct fell below the standard of care, and is in that sense wrongful, but it is not a tortious wrong. Rather, the wrong committed by the physician is an example of an unrealized wrong, and malpractice law, as a branch of tort, requires realized wrongs. Similarly, one who drives carelessly, taking his eyes off the road, may have acted carelessly, but not tortiously, if no person or property is harmed. It is an unreal-

⁴¹ See Model Penal Code § 211.2, at 194 (1980) (defining the crime of recklessly endangering another); see also *id.* § 211.2 cmt. 3, at 203 (noting that conduct need not actually endanger anyone to constitute the crime of reckless endangerment, so long as the actor does not actually know that no one is endangered).

ized wrong. Tort law does not impose any sanction on defendants under such circumstances, nor does it reward plaintiffs for pointing out or proving the occurrence of these wrongdoings.⁴²

The idea of realized wrongs can perhaps be better understood if we first distinguish between two ways of identifying actions for the purposes of classifying them as wrongs or not wrongs. Sometimes one can speak of actions in a manner that makes little or no reference to features outside of the actor's control. Suppose, for example, that speaker *S*, looking at a piece of jewelry, makes the following assertion: "That piece of jewelry contains a genuine emerald." We can answer the question of whether this speech act was a *sincere* assertion or an *insincere* assertion without answering the question of whether the jewelry does in fact contain a genuine emerald. Conversely, we cannot answer the question of whether *S*'s statement was a *true* assertion or a *false* assertion without knowing whether the jewelry contains a genuine emerald. We might, then, call the classifications of "sincere"/"insincere" as *externally independent* action classifications. For purposes of evaluating speech acts falling within these classifications, we need not know about facts external to *S*, such as the materials that make up the jewelry. All we need to do is gauge the sincerity of *S*'s beliefs. By contrast, the classifications of "true" assertion and "false" assertion constitute *externally dependent* action classifications: whether they are applicable depends on facts external to and independent of the agent's beliefs and conduct.

The question of whether a wrong has been realized is more akin to the question about the truth or falsity of *S*'s assertion than its sincerity or insincerity: It requires that we classify conduct in terms of externally dependent action classifications. Indeed, the very distinction between realized and unrealized wrongs is whether the defendant's potentially injurious conduct has in fact ripened into an injury. A defendant may act in a manner that risks an injury, or even in a manner that almost surely will bring about an injury, but that does not mean that the injury will actually occur. Note also that the classification of unrealized wrong is every bit as dependent

⁴² Dean Hurd has argued, along similar lines, that tort law embodies a conception of wrongdoing that requires risks to "materialize" before they can count as tortious wrongs. Heidi M. Hurd, *The Deontology of Negligence*, 76 B.U. L. Rev. 249, 262 (1996).

upon external facts as that of realized wrong, just as false assertion is no less dependent than true assertion upon external facts. If we do not know whether the jewel is an emerald, we do not know whether the utterance was true. If we do not know whether the injurious conduct actually ripened into injury, we do not know whether the conduct was realized.

This last point is potentially confusing, because there is a third category—the realizable wrong—that closely resembles a realized wrong, but that can be an externally independent action classification, just as sincere assertion can be. Driving in an unduly risky manner is, for example, a realizable wrong until such time as the wrongdoing ceases to have causal efficacy.⁴³ We can know whether someone has driven in a risky manner without knowing whether the risk has been realized. Whether some acts that threaten an as-yet unrealized physical harm may constitute a realized wrong (of a different type) will be discussed in Parts II–IV.

The category of inchoate crimes includes the class of unrealized and (at the time of prosecution) unrealizable wrongs. Attempt crimes, such as the failed effort described above in the bombing example, provide the most notable instances of this sort of wrong. Proof of the attempt crime may require proof that the wrong was at one time realizable; that there was an effort to murder, for example.⁴⁴ But it does not require that the wrong be realized and it does not matter that the wrong is no longer realizable.

The same is not true in torts. As noted earlier, a failed attempt to destroy someone's property, unattended by any other consequence for the victim, is not actionable in tort. And the same is obviously true for the other previously discussed examples. Many of

⁴³ Causal efficacy here invokes cause in both its factual and normative senses. The idea, in the famous language of Chief Judge Magruder, is that at a certain point the risks lurking within a wrongful act dissipate even though, as a matter of fact, the wrongful act ends up being a necessary step in a causal sequence leading to an injury unrelated to the risk. *Marshall v. Nugent*, 222 F.2d 604, 612 (1st Cir. 1955). Consider, for example, a speeding automobile driver who safely comes to a stop, then carefully opens his car door, in the process striking and injuring a passing cyclist. The wrong of speeding is not a proximate cause of the cyclist's injury precisely because the risks attending the misconduct (losing control, not having enough time to avoid a collision, etc.) had dissipated.

⁴⁴ We leave aside here issues of mistake and impossibility.

these actions can be sanctioned under the criminal law, but are not actionable in tort. The plaintiff must prove realization.⁴⁵

B. Tort as Civil Recourse

The non-actionability in tort of unrealized wrongs poses a significant problem for perhaps the dominant modern academic conception of tort law, which treats it as a system for deterring antisocial conduct. Whether they work within a loose utilitarian framework,⁴⁶ or a somewhat more technical economic theory,⁴⁷ most tort scholars have adopted the view that the main point of tort law is to provide legal incentives that will steer actors away from unjustifiably risky or socially harmful conduct by penalizing them if they engage in such conduct. If, however, tort law resembles criminal law in aiming to sanction socially harmful types of conduct, then we are left to ponder why tort law does not recognize inchoate wrongs. What, on the deterrence model, can explain this fundamental difference between tort and crime? The question becomes more acute when we move from criminal law to the regulatory law of the modern administrative state. On the deterrence view, safety regulations issued by agencies such as OSHA or EPA are even closer relatives to tort than criminal laws: They set standards, backed by fines or other sanctions that, in theory, will deter socially undesirable conduct. To the extent that regulations are enforced by public officials (as opposed to private citizens granted private rights of action), resultant harm is not always, perhaps not even typically, a threshold

⁴⁵ This is not to deny, of course, that many tort plaintiffs obtain settlements without proving the realization of a harm to them. See Goldberg & Zipursky, *Concern for Cause*, supra note 6, at 1415 n.13 (arguing that although the Agent Orange settlement allocated funds on a risk-of-harm basis, it did not announce or portend a new, risk-based approach to tort liability). Nor is it to deny that courts sometimes adjust burdens of proof for equitable reasons in a manner that permits plaintiffs to recover without having to prove that a particular defendant's tortious conduct caused injury to them. See Ripstein & Zipursky, supra note 6, at 231-44 (discussing burden-shifting in market share and alternative liability cases).

⁴⁶ See, e.g., Gary T. Schwartz, *Mixed Theories of Tort Law: Affirming Both Deterrence and Corrective Justice*, 75 Tex. L. Rev. 1801, 1816 (1997) (arguing that tort law is fundamentally concerned with deterring disutilitarian conduct, but that it pursues that goal within justice-based side constraints).

⁴⁷ See William M. Landes & Richard A. Posner, *The Economic Structure of Tort Law* 4-5 (1987) (describing the deterrence focus of utilitarian and economic theories of tort).

requirement for the issuance by an agency of regulatory sanctions against a wrongdoer. Why should it be a threshold requirement in tort law?

The absence of inchoate torts is not an anomaly, but points to a fundamental feature of tort law not fully captured by deterrence models. The differential treatment of inchoate wrongs in criminal law and tort law is explained by the different identity of the party bringing the action in each case. The state, representing the people, sometimes enjoys the power to punish for realizable wrongs and regulate for pure risk. Individuals (and the state when it acts in the capacity of an individual, for example, as a property owner suing for trespass) do not enjoy the power to demand the imposition of liability for any wrongs except realized wrongs. Of course, this just pushes the issue back a step. Now the question arises of how to explain the difference in treatment of suits prosecuted by the state versus those pursued by individuals.

Before sketching the answer to this question, a few preliminaries are in order. First, observe that our analysis provides a different focus than the standard deterrence-based account. The focus is no longer on the defendant, on how deserving of punishment the defendant is, or on the efficacy of sanctioning possible defendants. We are looking at the issue in terms of the identity of the party bringing the action. The question is why a tort plaintiff has rights that hinge on whether there is a realized wrong, but the state's powers to enact and enforce criminal and administrative law do not.

The second caveat is that the discussion thus far has contained generalizations that may suggest a certain naiveté on our part both about the distant history of tort and criminal law and about their current complexity. We are aware, however, of both historical and contemporary variations in criminal and tort law that permit, respectively, private actions for quasi-criminal punishment and public actions for private damages.⁴⁸ These do indeed complicate the

⁴⁸ See David Lieberman, *Mapping Criminal Law: Blackstone and the Categories of English Jurisprudence*, in *Law, Crime and English Society 1660-1840* (N. Landau ed., forthcoming 2002) (on file with the Virginia Law Review Association) (describing Blackstone's distinction between tort and crime in terms similar to the ones we draw, and noting that hybrid actions existed in his time); J. Randy Beck, *The False Claims Act and The English Eradication of Qui Tam Legislation*, 78 N.C. L. Rev. 539 (2000)

story, but their acknowledged existence *as hybrids* supports the basic themes of the account we are about to offer. Similarly, in insisting on the distinction between areas of law where injury is a requirement and those where it is not, we do not thereby mean to ignore that questions about what can count as a wrong or injury are often nuanced and often influenced by pragmatic considerations. Indeed, a central aim of Parts III and IV is to explore why the definition of actionable injury takes on the complex role that it does, and how the tort law goes about elaborating and revising its conception of injuries and wrongs.

So much for preliminaries. On our view, a plaintiff in a tort case has a right of action only because the defendant has committed a legal wrong against the plaintiff, that is, only if the plaintiff has suffered a wrong at the hands of the defendant. A right of action literally is a right, provided and mediated by our civil legal system, to act against the defendant in a particular way.⁴⁹ The reason the court system makes available rights of action in tort cases is that the system is built on the idea that those who have been wronged are entitled to some avenue of recourse against the wrongdoer. But, in a civil society, private violence is not permitted, even where there has been a legal wrong. The state therefore ordinarily must make some avenue of recourse available to the victim. It does this through the courts, via the tort system.

On our civil recourse conception, the black-letter rule that tort actions require an injury, and the precept *injuria absque damno* (wrong without harm), flow from this aspect of the normative structure of tort law. The point is not that defendants do not deserve liability unless their conduct amounts to a realized wrong. Nor is it that society would fail to benefit from their being sanctioned. The injury requirement, in other words, is not immediately concerned with whether a defendant ought to pay for her wrongs, either as a matter of punishment or deterrence. Rather, it is a requirement of standing set by tort law. When defendant *D* argues that plaintiff *P* has suffered no injury (and, hence, implicitly, that injury is necessary to *P*'s claim), *D* is arguing that *this* plaintiff

(discussing actions under the Federal False Claims Act in light of the history of *qui tam* actions).

⁴⁹ Benjamin C. Zipursky, Civil Recourse, Not Corrective Justice, 90 Geo. L.J. (forthcoming 2003).

ought not be permitted a right of action because, even granted that *D* acted wrongfully, *P* has not actually been wronged by her, and that *P* having been wronged by *D* is a prerequisite to being entitled to legal recourse against her.

This conception of recourse may strike some readers as archaic or barbaric because it links torts to vengeance or retaliation, a conception of tort that Holmes prominently and influentially rejected in *The Common Law* as inappropriate to modern society.⁵⁰ Perhaps, the objection might continue, it makes a certain amount of sense to say that only one who has been wronged by another can claim a right to retaliation, but that carries weight only if tort law is conceived as being about retaliation. Haven't we advanced beyond this barbaric conception of tort?

Clearly, society has advanced beyond clan feuds, duels, and other forms of officially sanctioned, private retaliatory violence. The idea of civil recourse is not an idea of private retaliation, however. On the contrary, some avenue of redress is required precisely because retaliation is no longer permitted. The courts settle disputes within a framework of civil law. But this is simply to say that our system of legal rights of action by putative victims against putative wrongdoers replaces a system of retaliation, not that it is one.

The false equation of recourse and retaliation ultimately rests on a confusion. It is one thing to say that, in a modern and civilized form of government, the right to private vengeance is eliminated and replaced by the operation of the state and its legal system. It hardly follows, however, that a modern and civilized form of government centralizes, or ought to centralize, all power to seek redress of wrongs through the state. The modern state arguably does maintain a monopoly on the business of imposing formal sanctions for legal wrongdoing, but that monopoly does not extend to prosecution—the power to trigger the apparatus of the legal system. The state, at least at a general level, retains the power to decide whether to enforce a demand for redress by the individual against another individual. And, more generally, the courts have created the rules that determine when such demands are to be enforced. But the power to make the demand, in private law, resides in the

⁵⁰ Holmes, *supra* note 2, at 2–5 (discussing the transition from ancient-barbaric to modern-civilized notions of justice).

individual. The principle found throughout the law of torts, that only a plaintiff who was injured by the defendant has a tort action against the defendant, derives from the basic normative structure of the private law. The power to sue rests on an entitlement to have some avenue of recourse against the defendant. In torts, this entitlement derives from the fact that the plaintiff has been wronged by the defendant, and yet is (almost always) barred by law (including criminal and tort law itself) from retaliation in the form of self-help.⁵¹

Criminal prosecutions operate in a different manner. The state (as prosecutor) has the power to demand that a defendant be punished. Again, the courts decide whether such a demand should be enforceable, here based on rules created by legislation. The criminal law, in contrast to tort law, does not contain a principle that prosecutions are only possible where a wrong has been perpetrated against a particular individual. That is because the state's power to punish does not entirely derive from an entitlement to redress a wrong to itself. The state's power to punish instead derives from its power to enforce publicly justifiable rules of conduct of a certain kind.⁵² A publicly justifiable rule of conduct may prohibit not only the doing of realized wrongs, but also the doing of even unrealized and unrealizable wrongs. It follows that a punishable violation of such a rule may exist even where a once-realizable wrong turned out to be unrealized.

The same may be said of the state's power to sanction the violation of legislative or regulatory safety standards. That power comes from the state's entitlement to lay down regulations for the public good and to enforce the regulations that it lays down.⁵³ Within the domain of legitimate regulations of conduct are ones that prohibit certain kinds of action, even where those actions are within the domain of externally independent action classifications. Hence, the state can require that one not drive above a certain speed, and it

⁵¹ Zipursky, *supra* note 49.

⁵² Here we refer to the "state" generically, not meaning to address issues of federalism. *United States v. Morrison*, 529 U.S. 598, 618 (2000) (noting that the American states' power to define and punish criminal conduct resides in their general "police power"); 4 William Blackstone, *Commentaries* *7 (noting the state's authority to enact criminal law "for the government and tranquility of the whole").

⁵³ See U.S. Const. art. I, § 8 cls. 1, 3 & 18.

can enforce these regulations. Indeed, inchoate crimes and regulations of risk-generating conduct-types are simply two different sub-categories of action classifications that are not dependent upon the occurrence of an injury.

On our account, then, there is a fundamental difference between tort law, on the one hand, and criminal and regulatory law enforcement, on the other. This difference goes to the question of where the power to enforce derives from. With respect to criminal and regulatory law, the power to enforce derives from the power to be the enforcer of legitimate rules in the public arena, which has been delegated to the state. The power to enforce in torts derives from having been the victim of a wrong. That is why a wrong must have been realized in the plaintiff, in tort, but no wrong needs to have been realized in criminal or regulatory law.

Today's most highly touted theories of tort law are, like Judge Posner's, regulatory models which treat the tort system as a privately enforced version of risk regulation.⁵⁴ It is merely useful, not fundamental, that tort actions are brought by parties who were wronged. It is useful both because the personal experience of having been injured lowers information costs and because the prospect of compensation for those injuries provides private parties with an incentive to enforce. Hence, tort law is an ancient (if misunderstood) version of what Congress has deliberately created in, for example, *qui tam* actions. As a bonus, according to the Posnerian account, tort law neutralizes what would otherwise be a socially counterproductive tendency of potentially injured parties to take too much caution in undertaking activities in which they might be injured if another party were to act negligently; the prospect of compensation through tort alleviates the need for inefficient levels of caution by potential plaintiffs.⁵⁵

Another commonly offered deterrence-based answer to this question of why tort requires realization appears equally unsatisfactory. The claim is that limiting tort to choate wrongs serves a second policy objective by allowing tort to compensate as it deters. This explanation is unsatisfactory for several reasons. Many tort

⁵⁴ See Goldberg, *supra* note 3 (manuscript at 36-38) (describing economic deterrence theory).

⁵⁵ Landes & Posner, *supra* note 47, at 62 (noting that the absence of driver liability in negligence to pedestrians may cause pedestrians to take excessive precautions).

claims are almost entirely non-compensatory, as for example are claims for punitive damages grounded in compensatory claims seeking only nominal or minimal damages. More fundamentally, ascribing a compensatory function to tort law does not by itself rule out a legal system that requires actors who commit unrealized wrongs to compensate for the real costs of injuries suffered by other actors engaging in similar but realized wrongs. The compensatory function of tort, in other words, could be met by any payment flowing from a wrongdoer to an innocent sufferer of injury: it does not by itself require payment to the victim by the person whose wrong injured the victim. Yet, tort law insists not only on the existence of wrongdoing and injury, but on the defendant's wrong having been realized in the form of an injury to the plaintiff.⁵⁶

The prevailing regulatory models are unsatisfactory because they fail to explain why a defendant's risky conduct should escape sanction in the absence of a loss, and why an uninjured plaintiff would not sometimes be an appropriate commencer of an action that aims to sanction risk-causing (or loss-causing) actors. The question becomes even more pressing when we realize that large areas of conduct generating peril to individual health and safety are not the subject of criminal law or more traditional regulation. Given that a good deal of risky conduct—driving carelessly, committing malpractice, failing to maintain one's premises safely—falls below the criminal and regulatory radar either formally or in practice, it might be socially desirable to have routine regulation of this conduct through civil causes of action that are not limited to suits brought by persons complaining of realized wrongs.

The standard Posnerian response to this objection is to argue that a system that permits actions for unrealized torts, while acceptable in principle, would overdeter in practice. As Part II will discuss, there are undoubtedly concerns, both pragmatic and principled, that counsel against the recognition of inchoate torts. But the threat-of-overdeterrence rationale hardly seems the most important. Courts already have means at their disposal to control for

⁵⁶ The importance of this sort of "bipolarity" to tort has been most forcefully emphasized by Professor Weinrib. See Ernest J. Weinrib, *The Idea of Private Law* 134–36 (1995).

overdeterrence. Indeed, as *Buckley* indicates, modern courts are not at all shy about invoking concerns over excessive liability as a ground for cutting off or capping liability. Why then draw the line at injury? Moreover, at least in some areas, the problem that we face today is more likely to be underdeterrence, considering difficulties in detecting certain forms of wrongful conduct, the inadequacy of damages to compensate injured plaintiffs fully, the existence of "cheap" settlements, the undercapitalization of corporate defendants, etc.⁵⁷ From a prescriptive point of view, these reasons suggest that a rationally designed deterrence-based conception of tort, particularly in an environment where a variety of risks are inadequately covered by direct regulation, would include liability for at least some classes of unrealized wrongs. From an interpretive point of view, they cast doubt on the plausibility of the concern for overdeterrence as an explanation of the injury requirement, which in turn casts doubt on the explanatory power of the deterrence model.

We also cast a skeptical eye on claims seeking recovery for mere risk exposure. But it is essential that the starting point be an understanding of the cases and an understanding of the reasons tort law is structured as it is. We therefore favor an interpretive framework that more plausibly explains what tort law actually is and does. The framework offered here better explains why it is that, in torts, a defendant will not be held liable for unrealized wrongs, but will be held liable for wrongs that are realized in an injury. Torts is not primarily a regulatory system, but a system in which private parties are afforded an avenue of redress against those who have wronged them.

To grasp the idea that civil recourse animates tort is not to uncover instructions for resolving a particular dispute. However, as we have argued elsewhere, the principles that undergird tort law at least provide a sort of default with respect to the practical question of adjudication.⁵⁸ In other words, it ought to take a significant combination of substantive moral, political, and economic arguments about the desirability of recognizing departures from the ideal of

⁵⁷ See, e.g., Richard L. Abel, *The Real Tort Crisis—Too Few Claims*, 48 Ohio St. L.J. 443 (1987) (arguing that tort claims are underlitigated by tort victims).

⁵⁸ See Goldberg & Zipursky, *The Place of Duty*, *supra* note 6, at 724–32.

civil recourse, as well as an institutional argument about the appropriateness of judges being the ones to undertake those departures, in order to justify substantial deviations from this framework.

Of course, if we shift perspectives, from adjudication to legislation, the interrelation of the normative and the practical should be characterized somewhat differently because legislatures, in our system, have greater discretion than judges to revise the law to serve desired goals. In the present context, the question becomes whether a legislator anxious to regulate certain forms of undesirable conduct should put her efforts into the particular project of legislatively modifying tort doctrine so that inchoate wrongs become actionable. It is far from obvious why this would be desirable. Such legislation typically would not be justified on the basis of the need of the victims—victims of realized wrongs are already empowered to sue, so the imagined legislation would only serve to empower those without injuries. Such beneficence, as Justice Breyer pointed out in *Buckley*, is not only unmotivated, it may actually cut against the recognition of liability for unrealized torts, since this will often supplant liability to those with ripened injuries.⁵⁹ The argument from a deterrence perspective, perhaps, is a better argument for modifying the common law to permit inchoate torts, but then it is hardly clear why a legislator would rationally select an agenda of private rights of action for persons who were put at risk but not otherwise harmed. If the risk creation problem is significant enough to capture legislative attention and warrant legislative action, why would a legislator think that the right response is to regulate the risk through tort plaintiffs as opposed to regulators and criminal prosecutors? At a minimum, one must consider an array of policy issues before concluding that legislative modification of the common law to recognize inchoate torts is the most promising path for regulation.⁶⁰

⁵⁹ *Buckley*, 521 U.S. at 443–44.

⁶⁰ See Goldberg & Zipursky, Concern for Cause, *supra* note 6, at 1419 (arguing that a proposed plan to allocate compensatory payments by gun manufacturers to shooting victims by reference to risk exposure gives inadequate consideration to a number of critical policy factors).

II. WHEN IS HEIGHTENED RISK A COGNIZABLE INJURY?

It is important to be clear about what the preceding argument has established, and what it has not. We have argued that, once tort law is properly conceptualized as a law of civil recourse, the notion of an inchoate tort becomes a self-contradiction. An inchoate wrong does not generate tort liability. Moreover, we have argued that the tort law's refusal to recognize inchoate wrongs, although somewhat mysterious on standard deterrence models, is quite intelligible on the civil recourse model. It is precisely because tort law is a law of civil recourse that plaintiffs may not invoke the machinery of the state to complain about wrongs that generate unrealized and unrealizable risks to them.

In arguing these points, it has *not* yet been established that enhanced-risk cases, such as suits complaining of heightened risk of cancer from exposure to a known carcinogen, must fail. To say that there are no inchoate torts is not to say that increased-risk claims are necessarily claims for inchoate torts. Indeed, such claims can be, and are typically, cast in a manner to indicate that a defendant's wrong has been realized, that is, that a plaintiff has suffered an injury.⁶¹ First, they may be cast to allege that the defendant's misconduct harmed the plaintiff simply by exposing the plaintiff to a still-extant risk of bodily harm or illness. Second, they may be cast to allege that the defendant's misconduct resulted in the *realization* of an injury, albeit not the illness risked by the exposure to the toxin. In short, the argument is that heightened risk is actionable because it is linked to another, realized injury, usually emotional distress or economic loss.

For reasons that will become clear, we refer to the harm associated solely with being subjected to a heightened risk of physical harm—a risk that is unaccompanied by any other realized injury—as an *intermediate* or unripened harm, in contrast to more traditionally recognized injuries as *ultimate* or target harms. Traumatic bodily harm and illness are the paradigmatic forms of ultimate harm. Emotional distress and economic loss, however, can also count as ultimate harms. This Part briefly analyzes attempts to frame claims for exposure to heightened risk of illness as claims

⁶¹ See *supra* text accompanying notes 11–12.

seeking compensation only for the intermediate harm of increased risk of illness.

Some have argued forcefully that, in and of itself, exposure to heightened risk of physical harm is not a harm to the person exposed.⁶² For purposes of this Article, however, we assume that exposure to increased risk can be regarded as in and of itself a loss of welfare to the person(s) placed at heightened risk. Whether correct or incorrect, this assumption strikes many as intuitive. For example, imagine that person *P* is given a choice, ex ante, between two states of affairs, *S*₁ and *S*₂. In *S*₁, *P* has a 2% risk of developing lung cancer. In *S*₂, *P* has a 10% risk of developing lung cancer. Otherwise, *S*₁ and *S*₂ are identical. Presumably, *P* will choose *S*₁. This example suggests to some that it is intelligible to talk about heightened risk of physical harm or illness as a harm to the person exposed to the heightened risk.

Nonetheless, even if we assume that heightened risk is a harm, it does not follow that this harm is an injury, that is, a cognizable harm.⁶³ In fact, the common law of negligence typically does not treat such harm as cognizable, as can be shown by offering modifications of the examples provided earlier in connection with the discussion of inchoate wrongs. Suppose a variation of *Vaughn v. Menlove*, in which Menlove has carelessly built the haystack on his property, increasing by 5% the likelihood that Vaughn's house will suffer fire damage as compared to the risk without the negligently constructed haystack. Even today, if Vaughn were to sue Menlove in negligence seeking compensation for the harm of heightened risk of property damage his suit would presumably fail. Unless the heightened risk ripens into the target harm of actual fire damage to person or property in the case of negligence, or the harm of interference with use and enjoyment of property recognized in the tort of nuisance—there is no cause of action. Of itself, the intermediate harm is not actionable.

The same reasoning applies to the shopping mall example. Suppose that mall owner *M* continues to operate his mall in a manner

⁶² E.g., Hurd, *supra* note 42, at 263–64; Stephen R. Perry, Risk, Harm and Responsibility, in *Philosophical Foundations of Tort Law* 321, 330–39 (David G. Owen ed., 1995).

⁶³ See Restatement (Second) of Torts § 902 cmt. a (1979) (explaining the injury/harm distinction).

that carelessly generates a heightened risk to patrons of being attacked by a third party (for example, because of poor lighting, the absence of adequate fencing, lack of security patrols, etc.). Even granted that each patron is in some sense harmed by the heightened risk of attack, until such time as a patron is attacked, there is no cause of action. These examples could readily be multiplied.

To say that the common law of tort does not treat pure heightened risk of physical harm as a cognizable harm by itself, of course, invites the question: "Why not?" The answer is part of a complicated story, too complicated to be analyzed in depth here. Essentially, we believe that the explanation lies in a proper conceptualization of the duty element of the negligence tort. The duty of care owed in most instances of actionable negligence is a duty to take care to avoid causing an ultimate harm, such as physical injury or property damage, not a duty to take care to avoid causing the intermediate harm of heightened risk. It is a duty to take care not to injure, rather than a duty to take care not to engage in injurious conduct (that is conduct that risks causing an ultimate injury).⁶⁴

This explanation, minimal as it is, will likely strike many readers as circular, but we do not think that it is. In our view, intermediate harms are not cognizable because the duties of care recognized by negligence law are almost always duties to avoid causing ultimate, not intermediate harms. Yet this argument invites a further question: Why does the common law of negligence rely on and enforce duties to avoid ultimate injuries rather than duties to avoid potentially injurious conduct? At least in theory, negligence could be built around duties to avoid causing risks of harm, although it has not. Likewise, negligence law could recognize a cause of action where the defendant breaches a duty not to cause a recognized ultimate injury even if the plaintiff can only demonstrate having suffered an unripened version of that injury. Yet it does neither of these things. Why not?

The answer to this question is over-determined by considerations of pragmatism and principle. First, as courts have noted, a system that permitted suits for intermediate harms would invite all

⁶⁴ Ripstein & Zipursky, *supra* note 6, at 217-21.

sorts of administrative headaches.⁶⁵ The volume of litigation would presumably increase substantially simply because many more persons are exposed to risk than suffer a realized risk of harm. So, too, redundant litigations would increase if persons were required first to litigate their increased risk claims, only to have to return to court years later to make out a claim when the risk is realized. Indeed, the concern for redundant and burdensome litigation underwrites the traditional rule against "claim splitting," which requires a plaintiff to bring all claims for injury caused by the same tortious conduct in one suit.⁶⁶ While courts have adjusted this rule in certain contexts, particularly in conjunction with medical monitoring claims,⁶⁷ the move to a general regime in which claims for enhanced risk are routinely separated out from claims of ultimate harm would seem to threaten a massive increase in administrative costs.

Second, until an ultimate harm has happened, decisionmakers will often have poor information about the risks in question, which means they will have a harder time than they already do in determining whether conduct should even be deemed careless, and, if so, how much compensation is owed to the plaintiff.⁶⁸ The experience of the fire damage or the parking lot attack, unfortunately, is often necessary to provide critical information regarding what precautions ought to be taken. This is not to deny that breach and damages pose problems of proof under the current system of liability for breaches of duties of non-injury. But the gravity of these problems would likely reach another order of magnitude were claims of intermediate harms rendered actionable.

⁶⁵ *Anderson v. W.R. Grace & Co.*, 628 F. Supp. 1219, 1232 (D. Mass. 1986) (noting that it would be unsound policy to encourage lawsuits in anticipation of grave disease) (citing *Gore v. Daniel O'Connell's Sons, Inc.*, 461 N.E.2d 256, 259 (Mass. App. Ct. 1984)).

⁶⁶ Kara L. McCall, Comment, Medical Monitoring Plaintiffs and Subsequent Claims for Disease, 66 U. Chi. L. Rev. 969 (1999) (noting the rule against claim splitting and arguing against its application to certain injuries).

⁶⁷ See *id.* at 983.

⁶⁸ See *Ayers v. Jackson Township*, 525 A.2d 287, 307 (N.J. 1987) ("A holding that recognizes a cause of action for unquantified enhanced risk claims . . . imposes on judges and juries the burden of assessing damages for the risk of potential disease, without clear guidelines to determine what level of compensation may be appropriate.").

Third, as Justice Breyer noted in *Buckley*, recognition of claims for intermediate harm creates serious prioritization concerns.⁶⁹ For any case in which the defendant has insufficient funds to pay all claims for intermediate and ultimate harms, it is likely that the former will frequently be compensated at the expense of the latter, simply because the former will arise earlier in time.

Fourth, and somewhat more speculatively, a system that recognized as actionable not only duties not to cause physical injury, but also duties not to cause various increments of risk of physical injury would potentially create undue de facto burdens on citizens' freedom of action. There is a sense in which each of us benefits from tort law being restricted to the enforcement of duties to take care not to cause ultimate harms. In this sense, tort law's notorious tolerance of so-called "moral luck"—the fact that, for reasons outside a wrongdoer's control, some wrongs ripen into injuries and result in liability whereas other identical wrongs do not ripen and do not result in liability—is not a tolerance for luck at all. Rather, it is a justified feature of a legal scheme concerned with recognizing liberty of action. It may be that, from certain moral perspectives, identical acts of careless driving, one of which ripens into a tort, the other of which does not, are equally deserving of sanction. Hence, from those perspectives, there is an element of arbitrariness as to when tort liability attaches.⁷⁰ But viewed from the standpoint of liberal political theory, the insistence on realization is not arbitrary. Rather, it harnesses chance to create a kind of buffer zone for free action: Unless and until injurious conduct actually causes an ultimate harm, it is not subject to sanction through a privately commenced lawsuit.⁷¹ In this regard, the employment of duties of non-injury in tort is akin to the rule against prior restraint of speech: It serves as a prophylactic by permitting a certain amount of undesirable conduct in order to ensure that liberty is preserved. Experien-

⁶⁹ *Buckley*, 521 U.S. at 442–44; see also Mark Geistfeld, *The Analytics of Duty: Medical Monitoring and Related Forms of Economic Loss*, 88 Va. L. Rev. 1921, 1942–43 (2002) (analyzing prioritization concerns).

⁷⁰ Christopher H. Schroeder, *Corrective Justice, Liability for Risks, and Tort Law*, 38 UCLA L. Rev. 143, 150 (1990).

⁷¹ Cf. Arthur Ripstein, *Equality, Responsibility, and the Law* 84 (1999) (suggesting that the problem of moral luck dissipates when tort is examined from a political-theory perspective); Ripstein & Zipursky, *supra* note 6, at 221–25, 229–31 (rebutting Professor Schroeder's critique of the injury requirement as arbitrary).

tially, this appears intuitive: Most people take advantage of the buffer zone created in part by the requirement of ultimate harm at one time or another, for example, by occasionally driving unreasonably. A system in which litigation could be instituted by anyone put at risk by allegedly unreasonable conduct would likely be seen and felt as burdensome, if not intolerable.

The latter suggestion about tort law's general refusal to recognize intermediate harms might be thought vulnerable to the following objection. As noted above, conduct of a harmful nature that does not cause an ultimate injury can be regulated through criminal and administrative law. In what way can it be a problem in liberal political theory to attach tort liability to such conduct? Again, however, this objection fails to appreciate the significance of the identity of the party seeking the sanction. Criminal and administrative law empower the government to place restraints on liberty by means of processes that, at least compared to the processes of tort law, tend to provide a certain degree of political accountability. This is most obvious in the case of democratically enacted legislation and notice-and-comment rulemaking, but it is also true to a certain extent of prosecutorial and regulatory decisionmaking. Prosecutors and administrators are informally accountable for their decisions to prosecute and, more significantly, face various legal and political constraints on how they exercise their power. By contrast, tort suits, particularly in a contingent fee system, are initiated by private actors who are unconstrained and unaccountable to anyone. Indeed, it is in many respects a great virtue of the tort system that a private plaintiff, at least, needs no one's permission to initiate a lawsuit. Moreover, she "prosecutes" her suit under legal standards less onerous than those of the criminal law. Precisely because it is so easy, in a political sense, to proceed with a claim in tort, a liberal regime will be cautious when arming individuals with the power to sue, lest suits become too prevalent and too invasive.⁷²

⁷² This point may help elucidate the emergence, within the Rehnquist era, of the constitutional jurisprudence of punitive damages. The current Justices tend to treat punitive damages primarily as a device for deterring future wrongful conduct through private lawsuits. Thus, to permit them is to permit regulation without ordinary political checks. To compensate for the absence of such checks, the Court has thus required beefed-up procedures after the initiation of the suit, as well as aggressive judicial review of the jury's damage award. See, e.g., *Cooper Indus. v. Leatherman Tool Group*, 532 U.S. 424, 431 (2001) (calling for *de novo* appellate review of juries' punitive dam-

Hence it is plausible to demand of tort plaintiffs, but not the government, that, ordinarily, they must await an ultimate injury before commencing suit.

Having offered these mostly suggestive remarks, we should also acknowledge the existence of what some might take to be exceptions to the rule requiring the plaintiff to have suffered an ultimate harm. Most significant are the so-called "loss of chance" cases.⁷³ As Professor Wex Malone noted fifty years ago, some courts early in the twentieth century began to relax the requirement of proof of causation in certain tort cases, including, for example, Jones Act cases involving suits by the survivors of drowned sailors claiming that a shipowner was negligent in failing to provide for, or undertake, a reasonable rescue effort.⁷⁴ Many of these incidents took place in stormy seas, and hours might have passed before the absence of the overboard sailor was noticed. For these and other reasons, plaintiffs had difficulty proving by a preponderance of the evidence that, had a given shipowner taken reasonable steps to rescue, a given sailor would have been saved.⁷⁵ In fact, the far more likely outcome was that no rescue would have occurred.

Faced with such cases, some courts began to reconceive the injury of which the sailors' survivors were complaining. The suits, they reasoned, were not seeking compensation for the death exactly, but for the sailors' "loss of a chance" of survival. The shipowner, as it were, was obligated to provide the sailors with whatever chance they had to live. Some courts expressed this idea in terms of a nontraditional duty to avoid causing intermediate harm. For example, the majority opinion in *Gardner v. National Bulk Carriers* reasoned that the shipowner owed to the sailor a "positive duty to make a sincere attempt at rescue."⁷⁶ On this account, "causation is proved if the [shipowner's] omission destroys the reasonable possibility of rescue."⁷⁷

age awards); *Honda Motor Co. v. Oberg*, 512 U.S. 415, 432-33 (1994) (expressing concern over the lack of procedural checks on punitive damages).

⁷³ See Dobbs, *supra* note 19, § 178, at 434-38.

⁷⁴ Wex S. Malone, *Ruminations on Cause-in-Fact*, 9 Stan. L. Rev. 60, 72-81 (1956).

⁷⁵ Dobbs, *supra* note 19, § 178, at 434-35 (describing the problem of proof to which the "loss of chance" doctrine is a response).

⁷⁶ 310 F.2d 284, 287 (3d Cir. 1962).

⁷⁷ *Id.*

In more recent years, the loss of a chance idea has been extended by some courts to the medical malpractice setting.⁷⁸ In the typical case, a physician negligently fails to diagnose a progressive disease, such as cancer, in her patient. Later, the patient learns of the condition, by which time the disease has worsened, lowering her chances of recovery. It turns out, however, that the odds of her surviving the disease were less than 50% to begin with, so that even if the patient had been properly diagnosed and treatment had been commenced earlier, she still likely would have died. Thus, if the patient dies, her surviving family members cannot bring a wrongful death claim, because they cannot prove that, had the doctor exercised due care, it is more likely than not that the patient would have lived. All that can be shown is that the doctor's carelessness reduced her chance of recovery by a certain amount (say, from 40% to 20%). Again, at least some of the courts that have permitted causes of action in similar scenarios have done so by reconceptualizing the doctor's duty from a duty to take care not to cause an ultimate injury (death) to a duty to take care not to diminish the plaintiff's chances of survival.⁷⁹ For example, in *Falcon v. Memorial Hospital*, the majority reasoned in part as follows:

We thus see the injury resulting from medical malpractice as not only, or necessarily, physical harm, but also as including the loss of opportunity of avoiding physical harm. A patient goes to a physician precisely to improve his opportunities of avoiding, ameliorating, or reducing physical harm and pain and suffering.⁸⁰

If loss of chance cases such as *Gardner* and *Falcon* could be fairly characterized as the "camel's nose" for a revolution in tort law under which duties of non-injury were transformed, across the board, into duties of non-injuriousness, then this Part's argument might be in jeopardy. That has not been the case, however, and this account may help explain why.⁸¹ Here, it is vital to realize that even

⁷⁸ See Dobbs, *supra* note 19, § 178, at 434–35.

⁷⁹ *Id.*

⁸⁰ 462 N.W.2d 44, 52 (Mich. 1990), superseded by statute, as stated in *Weymers v. Khera*, 563 N.W.2d 647, 652–53 (Mich. 1997) (en banc).

⁸¹ See Dobbs, *supra* note 19, § 178, at 435 (noting decisions refusing to recognize loss of chance claims where causation cannot be proven by a preponderance of the evidence).

those courts that have recognized loss of chance claims have only done so if two conditions are present. First, loss of chance has always been recognized as a second-best solution for dealing with claims by persons who have in fact suffered a physical injury or death. In other words, loss-of-chance theories have been developed in response to perceived inequities faced by certain tort plaintiffs who have demonstrably suffered ultimate injuries, but face a particular sort of difficulty in proving causation. Put simply, there is no occasion to engage in loss-of-chance analysis unless a class of injured plaintiffs faces a systematic inability to prove that the negligent acts of members of an identifiable class of defendants caused them injury. The recasting of the duty in these cases as a duty of non-injuriousness rather than non-injury only occurs because, in the absence of the recasting, an entire class of plaintiffs will be denied all recovery, and an entire class of wrongdoers, some of whose wrongdoings ripened into injury, will be left free of all liability. In this respect, loss of chance resembles the device of market share liability, by which the courts shift the burden of proof of causation because there is no information as to which among multiple manufacturers caused a particular plaintiff's injuries.⁸²

Second, the epistemic problem (systemic absence of proof of causation) is a necessary, but not sufficient, condition for reconceptualizing the duty in the loss-of-chance cases.⁸³ In addition, courts have required that the original duty of non-injury contain or support something in the nature of an affirmative duty to assist or protect, rather than a purely negative duty to refrain from engaging in risky conduct. Where the original duty of non-injury includes affirmative-duty components, it makes particular sense, in light of the proof problems discussed above, to treat loss of opportunity as a cognizable harm. Again, to quote the *Falcon* opinion:

A woman who engages the services of a physician . . . does so to reduce pain and suffering and to increase the likelihood of her surviving and the child surviving childbirth in a good state of health even though the likelihood of the woman and child not surviving in good health without such services is far less than fifty

⁸² See Ripstein & Zipursky, *supra* note 6, at 231–37.

⁸³ Dobbs, *supra* note 19, § 179, at 438–39 (noting the limited contexts in which duty is imposed).

percent. That is why women go to physicians. That is what physicians undertake to do. That is what they are paid for.⁸⁴

In other words, it is plausible to characterize the doctor-patient relationship as including an affirmative duty on the part of the doctor to take steps to give her patient every chance to recover, not just to take care to avoid causing those injuries.⁸⁵ Ordinarily, this aspect of the duty of care owed by a doctor to his patient is merely implicit within an allegation of malpractice. If, for example, an OB-GYN neglects to follow reasonable procedures, such as failing to determine the Rh status of mother and baby, a claim alleging that an ultimate harm resulted from the malpractice will not rely upon or highlight this affirmative aspect of the physician's duty. In other words, a failure to follow basic standards of reasonable care is always a breach of the stronger affirmative duty to take care to maximize the chances of a good outcome. Sometimes, however, if it is impossible for the plaintiff to establish a breach of the duty to take care not to cause physical injury, the affirmative nature of the physician's duty emerges or is highlighted. Its residual component is revealed, supporting a claim for breach of the duty to provide the patient with the best chance for survival. By contrast, for example, Menlove's duty to his neighbor Vaughn, being strictly a negative duty, contains no such residual affirmative duty. Thus if Vaughn somehow could not prove that Menlove's negligence probably caused his house to burn down, Vaughn likely would not be able to resort to the second-best argument that Menlove had injured him by increasing the risk that his house would burn down.

The foregoing analysis of loss-of-chance cases is admittedly suggestive rather than comprehensive. It is offered only to show that, in recognizing pockets of liability for loss of chance, courts have not committed themselves (and need not commit themselves) to the recognition of "risk" as a cognizable harm. Recovery for loss of chance, in other words, is not recovery for heightened risk, rather it

⁸⁴ *Falcon*, 462 N.W.2d at 52.

⁸⁵ See Goldberg & Zipursky, *Concern for Cause*, *supra* note 6, at 1416-17. In our view, then, it is probably incorrect to describe the loss-of-chance cases as permitting recovery for increased risk of harm. Professor Perry has argued that they permit recovery for loss of an opportunity to obtain adequate treatment. See Perry, *supra* note 62, at 339.

is recovery for breach of a special kind of affirmative duty, on the occasion of a full-fledged physical injury.

III. RISK OF FUTURE INJURY AND THE LAW OF EMOTIONAL DISTRESS

Thus far we have argued that claims for future injury are not well framed in negligence either as claims for inchoate wrongs, or as claims for the harm of having been exposed to risk of physical injury. Although this discussion has been necessary to clarify the issues at stake, it likely will read to many as an extended prelude. As noted in the introduction, lawyers and judges tend not to analyze claims for heightened risk along the lines thus far discussed, although they might unwittingly slip into such analysis. Rather, they are much more likely to analyze such claims under the categories of negligent infliction of emotional distress ("NIED") and negligence causing pure economic loss. We now turn to examine the former category, commencing with a brief review of the state of current doctrine.

A. The Current State of Emotional Distress Doctrine

1. Basic Doctrine

Although it is regularly employed in ordinary academic discourse, the use of NIED as a category of negligence liability has always been awkward. The last century did, of course, witness important expansions of liability in cases where the primary compensable injury consisted of a plaintiff's emotional distress.⁸⁶ Nonetheless, analysis of this area is rife with confusion.

The most fundamental confusions arise from the tendency among scholars and judges to conceive of NIED as permitting recovery for "pure" emotional distress. To be sure, this usage is helpful in that it starkly contrasts NIED claims from cases in which the plaintiff is compensated for pain and suffering and other forms of distress that attend a physical harm caused by defendant's carelessness. The latter sort of case is by far the more common case in

⁸⁶ Dobbs, *supra* note 19, § 308, at 836.

which emotional distress is compensated.⁸⁷ To take a contemporary example, plaintiffs who are able to prove that a defendant's negligence in exposing them to asbestos caused them to develop mesothelioma are typically permitted to recover for fears concerning their future health as part of the general emotional distress associated with having mesothelioma. This is a special case of the more general rule holding that emotional distress is not generally sufficient to count as the predicate injury in a negligence case, but is compensable if parasitic upon a proven physical injury.

The problem with referring to suits that seek compensation for emotional distress that is not parasitic on a physical harm as claims for "pure" emotional distress is that the term "pure" connotes the absence not just of a predicate physical harm, but of any other harm—including *consequent* physical harm, as well as predicate *intangible* harm, such as the harm of being deprived of one's liberty of movement. It further suggests that that emotional distress, like physical harm, should be conceived of as a harm unmediated by the immediate circumstances giving rise to the distress.⁸⁸ Both of these connotations are unfortunate because, under current law, cognizable claims for emotional distress that are not parasitic on a physical harm are rarely "pure" in these ways. The phrase "pure emotional distress" suggests an orbit of duty and liability that is much broader than any recognized under current law. It also fails to capture what is most salient about the situations in which courts do permit such claims to go forward.

Consider the so-called "fright" or "zone of danger" cases, mentioned in the introductory discussion of the Supreme Court's deci-

⁸⁷ Id. § 302, at 822 ("Far and away the most common claims of emotional distress... are those associated with personal injury actions.").

⁸⁸ Physical harm is "unmediated" in that the immediate circumstances giving rise to the physical harm are always the same—the interaction of some force or agent with the plaintiff's body. Thus, if *B* acts carelessly toward *A* so as to proximately cause *A* to suffer a broken leg, it doesn't matter whether the break was caused because *B* knocked *A* over, committed malpractice on *A*, or furnished *A* with a bone-dissolving drug. By contrast, as we explain below, when *A* sues for negligence causing emotional distress, it matters very much to *A*'s chances of recovery how that distress was induced. In many jurisdictions, for example, absent a special relationship between *A* and *B*, if *A*'s emotional distress did not result from *B*'s carelessness causing *A* to apprehend imminent harmful bodily contact, *A* cannot recover from *B*.

sion in *Consolidated Rail Corp. v. Gottshall*.⁸⁹ To understand the scope of the fright doctrine, one must first appreciate two older doctrines it was developed to supplant. The first served as a broad bar to emotional distress claims and provided that a physical injury or illness brought about by fear, anxiety, or shock could never count as having been proximately caused by a defendant's negligent act.⁹⁰ As a matter of law, the connection between the negligence and injury was deemed in these cases to be too attenuated. The second doctrine, which was designed in some ways to blunt the force of the proximate cause rule, held that *any* impact or bodily touching, no matter how minimal, would count as a predicate physical injury.⁹¹ Under this pair of rules, which was in place in the late nineteenth and early twentieth centuries, a plaintiff who was subject to the least touching as a result of the defendant's negligence could recover even if all of her claimed damages were for emotional distress, whereas a person subjected to a traumatic episode in which she was not touched recovered nothing.⁹²

Dissatisfied with this exercise in line-drawing, courts in the mid-twentieth century came to recognize a cause of action for negligent infliction of physical injury through fright. To make out such a claim, the plaintiff is typically required to prove that the defendant's negligence (a) caused the plaintiff to apprehend that she was in danger of imminent bodily harm, and (b) that this apprehension in turn manifested itself in a physical injury such as a miscarriage

⁸⁹ See *supra* text accompanying notes 16–21 (discussing the Court's application of *Gottshall* in *Buckley*).

⁹⁰ The classic analysis and critique of this proximate cause argument is Frances H. Bohlen, *Right to Recover for Injury Resulting from Negligence Without Impact*, 50 Am. L. Reg. 141 (1902), *reprinted in* Frances H. Bohlen, *Studies in the Law of Torts* 252–90 (1926); see also *Battala v. State*, 176 N.E.2d 729, 731–32 (N.Y. 1961) (criticizing the public policy reasons given for denying no-impact emotional harm claims).

⁹¹ See William L. Prosser, *Handbook of the Law of Torts* § 34, at 214 (1941).

⁹² The poster child for the inaptness of the regime created by combining the proximate cause and impact rules was *Mitchell v. Rochester Railway Co.*, 45 N.E. 354, 355 (N.Y. 1896), overruled by *Battala v. State*, 176 N.E.2d 729, 731–32 (N.Y. 1961), which denied recovery to a woman who suffered a miscarriage allegedly as a result of the serious fright she received when defendant's runaway horse nearly ran her over, but did not actually touch her. See Prosser, *supra* note 91, § 34, at 214–15 (discussing *Mitchell*'s "untenable notions of causal connection" while lampooning the proximate cause and impact rules).

or diagnosable illness.⁹³ As the Second Restatement emphasizes, courts permitting recovery in the fright cases have generally not taken themselves to be recognizing a cause of action for "pure" emotional distress.⁹⁴ Instead, they permit claims for physical illnesses that are generated through the particular mental medium of fright. In this respect, the fright causes of action were primarily a repudiation of the old proximate cause doctrine. This repudiation, it was hoped, that would largely obviate the need to rely on the crude counterweight of the impact rule. *Buckley* itself confirms this understanding of the fright cases by ruling, in effect, that, once *Gottshall* recognized a cause of action for fright arising under FELA, there was no longer a need to permit claims based on not-presently-harmful impact with asbestos dust.⁹⁵ Instead, a plaintiff must either seek damages for emotional distress parasitic on a genuine physical injury or illness, or he must make out a claim for negligent infliction of physical harm through fright.⁹⁶

In the predicate physical injury and fright cases, which have, until recently, made up the vast bulk of cases in which a victim of negligence has received compensation for emotional distress, the distress is compensable, but not in itself actionable. Nevertheless, courts do purport to recognize two pockets of liability for "pure" emotional distress. First, there are the so-called "bystander" cases. These permit recovery for emotional injury to the plaintiff that occurs when a plaintiff observes firsthand the defendant negligently injuring her spouse, child, or other close relative.⁹⁷ As originally conceived, these bystander claims, too, appear to have been understood as claims for *physical* injury or illness caused through the medium of nervous shock.⁹⁸ However, courts have more recently

⁹³ See, e.g., *Amaya v. Home Ice, Fuel & Supply Co.*, 379 P.2d 513, 515 (Cal. 1963) (noting that fright causes of action must include allegation of "consequent bodily illness"), overruled on other grounds, *Dillon v. Legg*, 441 P.2d 912, 925 (Cal. 1968).

⁹⁴ Restatement (Second) of Torts § 313(1) (1965) (noting that the action is for emotional distress resulting in illness or bodily harm); id. § 313(1) cmt. a ("[T]his section does not give protection to mental and emotional tranquility in itself.").

⁹⁵ See *Buckley*, 521 U.S. at 429–30 (rejecting the impact associated with dust inhalation as sufficient to support an award of emotional distress damages).

⁹⁶ See *supra* text accompanying notes 12–21 (discussing *Buckley*'s rejection of the plaintiff's NIED claim).

⁹⁷ *Dobbs*, *supra* note 19, § 309, at 839–41.

⁹⁸ E.g., *Dillon v. Legg*, 441 P.2d 912, 920 (Cal. 1968) (permitting bystander recovery, but confining its holding to instances of wrongful conduct causing physical injury).

suggested that they are in fact claims for "pure" emotional distress.⁹⁹ Essentially, they allow relatives standing just outside the zone of danger to recover for their distress, even though they were not concerned for their own physical well being.

Second, liability attaches for emotional distress when the negligent conduct occurs within the context of certain "special relationships." The key factor is that the situation in which the parties interact attests that the defendant implicitly undertook to be vigilant of the emotional well-being of the plaintiff.¹⁰⁰ Thus, doctors and therapists have been held liable for negligently distressing their patients.¹⁰¹ Likewise, morticians have been held liable for upsetting family members by mishandling of the decedent's corpse.¹⁰² In older cases, telegraph companies were held liable for being careless in falsely informing the recipient of a family member's death, or in failing to deliver a death telegram to the decedent's relative, thereby denying him or her the opportunity to attend the decedent's funeral.¹⁰³

2. *Fear-of-Disease Doctrine*

The fear-of-disease cases have emerged in the last fifteen years against this doctrinal backdrop. In many jurisdictions, they have caused something of a throwback as courts have once again begun to invoke the old "impact" rule. Like their nineteenth-century predecessors, these courts do so as a way of simultaneously deny-

⁹⁹ E.g., *Thing v. LaChusa*, 771 P.2d 814, 829-30 (Cal. 1989). Even thus cast, one may doubt that these are really actions for the wrong of carelessly causing another to suffer emotional distress. Indeed, in our view, bystander claims are best understood as judicial extensions of "derivative" actions for wrongful death and loss of consortium, whereby a plaintiff in a familial relationship with the victim of another's negligence is granted a special dispensation to sue as the vicarious beneficiary of a wrong causing physical harm to the victim. See Benjamin C. Zipursky, *Rights, Wrongs and Recourse in the Law of Torts*, 51 Vand. L. Rev. 1, 28-29 (1998) (suggesting that courts in emotional distress cases sometimes grant "standing" to those who have not been wronged by the defendant).

¹⁰⁰ Dobbs, *supra* note 19, § 308, at 836-37.

¹⁰¹ *Id.* § 312, at 849-50.

¹⁰² *Id.* § 308, at 836-37.

¹⁰³ *Id.* In our view, these special relationship cases are perhaps the only true instances of liability for "pure" emotional distress: This sort of plaintiff fairly complains that the defendant breached a duty to be vigilant of the plaintiff's interest in emotional well-being *tout court*.

ing "pure" NIED claims while watering down the predicate physical injury requirement, thus permitting readier recovery of emotional distress damages as "parasitic" on a mere touching. In some states, for example, a plaintiff can recover for fear of cancer if she can prove that the defendant's negligent conduct caused a physical impact upon her, in addition to causing her fear of cancer.¹⁰⁴

A second variation, closely related to the revival of the impact rule, involves conferring legal significance on subcellular changes or changes to the immune system. Here, the physical changes are neither the predicate injury nor the consequent injury, but rather the physical realization of the risk factors caused by defendant's negligence. While these physical changes themselves do not diminish plaintiff's physical functioning or cause pain or discomfort, they are deemed cognizable harms of a sort that will support recovery for emotional distress over the fear of future illness.¹⁰⁵

Other jurisdictions have taken a rather different approach, requiring a consequent physical symptom or manifestation of emotional harm (such as fear or fright) in order to recover for the fear itself. These differ from the "fright" cases, in that they do not conceive of the plaintiff as seeking compensation for physical injury consequent on emotional trauma. They also differ from the predicate injury cases in that the tort is not thought of as the negligent causing of the physical injury, but as the negligent causing of fear of disease. Instead, the symptom or manifestation serves an evidentiary role, one that attests to the sincerity and gravity of the claim for emotional distress damages.¹⁰⁶

Still other courts hold that a plaintiff cannot recover for fear of disease without proof that she was actually exposed to the carcinogenic or pathogenic agent (as opposed to merely having been touched, which would trigger liability under the impact rule). This variation is sometimes applied in fear-of-cancer or fear-of-hepatitis cases, but it is particularly common in fear-of-AIDS cases. Indeed,

¹⁰⁴ *Adams v. Clean Air Sys.*, 586 N.E.2d 940, 942 (Ind. Ct. App. 1992); see also *Sanders v. Brown*, 343 S.E.2d 722 (Ga. Ct. App. 1986) (stating the impact rule in a pesticide exposure case).

¹⁰⁵ See, e.g., *Anderson v. W.R. Grace & Co.*, 628 F. Supp. 1219, 1226-27 (D. Mass. 1986) (holding that subcellular changes are adequate to support a NIED claim).

¹⁰⁶ *First Nat'l Bank v. Drier*, 574 N.W.2d 597, 600 (S.D. 1998) (requiring objective symptomology).

fear-of-AIDS plaintiffs are typically required to prove that they were exposed to HIV under conditions that would have made its transmission possible.¹⁰⁷

A fifth and less physically focused rule is found in the California Supreme Court's prominent opinion in *Potter v. Firestone Tire and Rubber Co.*¹⁰⁸ Plaintiffs in that case argued that they should be able to recover from Firestone, which had regularly disposed of toxic waste in a landfill near their homes, causing them a substantially increased risk of cancer.¹⁰⁹ Firestone argued that since there was no proven physical injury, there could be no recovery.¹¹⁰ The California Supreme Court held that plaintiffs could recover in negligence for fear of cancer without predicate physical injury or physical impact, but only if they could prove that it was likelier than not that they would develop the disease.¹¹¹ The rationale for adapting what appeared to be a test for the cause-in-fact requirement was, in the first instance, the observation that fear of illness ought to be reasonable before being compensable. Of course, the court also adverted to the need to control a potential flood of litigation, as well as the need for objective criteria to prevent fraud.¹¹²

Finally, a small number of courts have rejected bright-line rules entirely in this area, and instead adopted standards aimed to evaluate directly the strength of each plaintiff's claims. Hence, for example, some courts will look for strong evidence of the seriousness of the emotional distress; some courts will look hard at the reasonableness of the particular plaintiff's fear; some courts will ask—as in traditional intentional infliction of emotional distress cases—whether the distress in question was more than a normally constituted person could be expected to endure.¹¹³

¹⁰⁷ Dobbs, *supra* note 19, § 311, at 845 (noting that courts presented with fear-of-AIDS claims "usually" require evidence of exposure).

¹⁰⁸ 863 P.2d 795 (Cal. 1993).

¹⁰⁹ *Id.* at 801.

¹¹⁰ *Id.* at 803.

¹¹¹ *Id.* at 816. The plaintiffs had not met the more likely than not standard, but the court remanded the case for a determination of whether Firestone could be held liable on the alternative theory of intentional infliction of emotional distress. *Id.* at 827.

¹¹² *Id.* at 816.

¹¹³ See, e.g., *Williams v. Warren Gen. Hosp.*, 684 N.E.2d 730, 733 (Ohio Ct. App. 1996) (permitting recovery where distress is severe and debilitating).

As numerous courts and scholars have indicated (and this conference attests), the law of "fear of disease" and "fear of injury" warrants close evaluation and thoughtful development. The courts are casting about for a sensible way of framing these claims, and have adopted divergent approaches. Moreover, many of the standards currently in use cry out for scrutiny and skepticism. The revival of the physical impact rule, for example, is a glaring illustration of doctrine that is forced and formalistic. Likewise, the *Potter* court's like-her-than-not rule is intuitively unsound. More generally, these cases raise basic questions about negligence law. Does the basic distinction between predicate injury and parasitic damages make sense? How ought courts handle the problems of hypersensitive plaintiffs, too much litigation, and the threat of fraudulent claims?

This Part first criticizes the ways in which many scholars and courts conceptualize fear of injury and emotional distress claims more generally. The basic point is that courts use concepts that are insufficiently robust and nuanced to permit the emergence of a coherent doctrine with respect to claims for negligence causing emotional distress. In their place, we propose an alternative way of conceptualizing these cases in terms of a duty to take care to avoid placing plaintiffs in certain kinds of threatening situations. Finally, we offer some reasons for thinking that this understanding will be useful to courts faced with the practical question of how to dispose of fear-of-injury cases.

B. The Current Conceptualization of Emotional Distress Claims

As previously discussed, many of the lines drawn by the courts seem at first blush arbitrary and indefensible. Why must a fear-of-AIDS claimant prove actual exposure to HIV if, for example, she has reasonable grounds for such fear? Why, in *Potter*, did the California Supreme Court require a fear-of-cancer claimant to prove that his fear was grounded in a "more likely than not" probability that he will develop the disease? The trouble stems in part from the impoverished conceptual framework employed by most scholars and many courts to analyze negligence claims. This framework is neither troubled by, nor helpful in overcoming, these instances of arbitrariness. Indeed, for reasons that we now explore, the framework suggests that courts have no choice but to draw arbitrary lines

in the area of negligent infliction of emotional distress. On this account, about the most one can hope for in this area is that the courts will be candid in acknowledging the arbitrariness of their rules, as the California Supreme Court sometimes has been.¹¹⁴

The standard conceptualization starts from an intuitive premise. Indeed, its intuitiveness is precisely why the framework has become standard. We label it the "equivalence-of-harms" thesis (the "equivalence" thesis, for short). It asserts that, to the victim who experiences the harm, the type of harm—as opposed to its magnitude or severity—is irrelevant.¹¹⁵ Assuming each of several types of harm on the same order of magnitude, there is no normative distinction to be drawn between them. Insofar as tort law is concerned, there is no essential difference between traumatic physical injury, illness, or psychic distress. Simply put, harm is harm is harm.¹¹⁶

Under the equivalence thesis, tort law generally, and negligence law in particular, appear facially odd and thus badly in need of explanation or reform. Both have long divided up the liability universe along the lines rejected as incoherent by the equivalence thesis. As previously seen, courts have long given "second class" citizenship to emotional distress and intangible economic loss as harms or protected interests.¹¹⁷ Moreover, as the bevy of special rules invoked by courts in the fear of diseases cases attests, they have continued to do so even as they have expanded liability for negligence causing emotional distress.

The ill fit between negligence doctrine and the equivalence thesis invites two obvious questions: How could courts have gone down this road in the first place and why do they still insist, to some degree, on staying that course? The now-standard answer to the first question is that the nineteenth-century judges who first

¹¹⁴ E.g., *Thing v. LaChusa*, 771 P.2d 814, 828 (Cal. 1989).

¹¹⁵ Prosser, *supra* note 91, § 34, at 211–12.

¹¹⁶ Often in discussions about tort liability for emotional distress, the equivalence thesis is connected to a claim about neuroscience. It is only a matter of time, the claim goes, until the line between psychic and physical injuries collapses completely in the face of mounting knowledge about how the physical operations of the brain link up to emotional states. At that point, the equivalence thesis will be established not merely as a matter of theory or experience, but as hard science. See *id.* § 11, at 55–56.

¹¹⁷ See *supra* note 19 (citing Professor Dobbs' discussion of courts' traditional reluctance to recognize NIED claims).

elaborated the modern law of negligence labored under various pre-modern misconceptions about science and nature that led them to discount emotional injury as insubstantial or unreal.¹¹⁸ Moreover, their illusions were tied to a generally sexist disposition. Emotional distress was, to their eyes, the sort of thing that women, in all their delicacy, experienced.¹¹⁹ As women's interests were, at the time, of little concern to the law, emotional distress was not either.

Assuming its validity, this explanation still leaves the second puzzle. On the whole, modern judges no longer operate under either the strongly sexist conception that women's interests are of no legal concern, or that women are emotionally delicate in a way that men are not. Moreover, judges presumably no longer employ nineteenth-century, pseudo-scientific notions. Yet, as we have seen, they still show a general reluctance to push negligence law where it ought to be under the equivalence thesis. For negligent conduct (as opposed to nonfeasance) that causes *physical harm*, the general rule is that the defendant is on the hook for any foreseeable harms proximately caused by his negligent conduct.¹²⁰ Indeed, as told by influential scholars such as the late Professor Gary Schwartz, the great "story" of negligence law in the twentieth century—as it applies to physical harms—was the casting off of the various archaic and arbitrary limits on its reach, such as the privity rule, special duty rules for landowners, charitable immunity, contributory negligence, and so on.¹²¹ With this, one witnessed the emergence of a "pure" fault principle under which liability attaches, *prima facie*, to any unreasonable conduct that in fact causes another a physical harm.¹²² Yet Professor Schwartz never supposed that courts had ar-

¹¹⁸ See Prosser, *supra* note 91, § 11, at 54–57 (noting and debunking various arguments for denying the legally compensable character of emotional distress injuries).

¹¹⁹ *Id.* § 11, at 55 ("It is not difficult to discover . . . a distinctly masculine astonishment that any woman should ever be so silly as to allow herself to be frightened or shocked into a miscarriage.").

¹²⁰ Dobbs, *supra* note 19, § 227, at 578.

¹²¹ Gary T. Schwartz, *The Vitality of Negligence and the Ethics of Strict Liability*, 15 Ga. L. Rev. 963, 963–68 (1981); see also Gary T. Schwartz, *The Beginning and the Possible End of The Rise of Modern American Tort Law*, 26 Ga. L. Rev. 601, 605–09 (1992) (exploring why judges were motivated to expand negligence liability to its full extent).

¹²² See Goldberg & Zipursky, *The Place of Duty*, *supra* note 6, at 660, 664 (noting that the Third Restatement, as drafted by Professor Schwartz, contains this formulation of the *prima facie* case of negligence causing physical harms).

rived at a comparable general principle of liability for NIED. Instead, we see only the partial and fragmented patchwork of reforms noted above. Moreover, in place of the older formalisms, modern courts have introduced formalisms of their own. Indeed, the standard-bearer of instrumentalism, the California Supreme Court, is among the worst offenders here, as evidenced by its insistence that only close relatives who contemporaneously observe the injury may recover under the theory of bystander liability,¹²³ and its insistence on the likelier-than-not requirement in *Potter*.¹²⁴ The most emphatic recognition of this divergence between the law of negligent infliction of physical harm and NIED is found in the basic negligence provisions of Professor Schwartz's draft of the Third Restatement,¹²⁵ which carve out emotional distress claims for separate treatment from claims for physical harm on precisely this ground.¹²⁶

What, on the standard view, explains these continued departures from the equivalence thesis? The answer, its adherents suppose, is to be found in the interaction between principle and practice. In principle, the law of negligence ought to incorporate the equivalence thesis: The law ought to contain a pure fault regime that attaches liability to any person's negligent conduct that in fact causes any sort of harm to another. In practice, however, it simply cannot. There is too much distress in the world traceable to the negligent acts of others to suppose that it should be *prima facie* actionable. Moreover, too much of this distress takes the form of transitory upset, rather than lasting harm warranting compensation and deterrence. Finally, and probably most importantly, claims of emotional distress are, as contrasted to claims of physical injury, harder to substantiate, easier to fake, and harder to value. For all of these reasons, the operation of a pure fault regime that incorporates the equivalence thesis would in practice be too prone to error, abuse,

¹²³ *Thing v. LaChusa*, 771 P.2d 814, 828–29 (Cal. 1989) (noting that denial of recovery to close friends and partners is arbitrary, yet necessary to avoid floodgates and fraud problems).

¹²⁴ See *supra* text accompanying notes 108–112 (describing the holding in *Potter*).

¹²⁵ See Goldberg & Zipursky, *The Place of Duty*, *supra* note 6, at 664 (quoting the relevant commentary from the Draft Restatement).

¹²⁶ Martha Chamallas, *Removing Emotional Harm from the Core of Tort Law*, 54 Vand. L. Rev. 751, 753 (2001).

overdeterrence, and overcompensation.¹²⁷ And so the courts have settled on various limitations on claims for negligent infliction of emotional distress, including the predicate injury/parasitic damage distinction, the requirement that fright plaintiffs prove a physical illness, the limitation of bystander liability to close relatives who observe the injury, and so on. The difference is that these limits, unlike the old proximate cause limit, are employed not for reasons of faulty metaphysics or sexism, but because of sound considerations of policy.

Unfortunately, although these lines are now supposedly well-grounded, the policies on which they rest tend not to contain within them guidance as to how these lines might be drawn in a principled manner. Perhaps some of the foregoing policy considerations contain such guidance. To reduce trivial and fraudulent claims, for example, it might make sense for courts to insist that emotional distress plaintiffs provide evidence from a treating physician or other competent expert as to the existence of his or her distress and its severity. (That is, assuming there is not a cadre of experts prepared to testify to distress at the drop of a hat.) Other objectives—particularly the concern about too much liability—offer no principled basis for deciding where to cut off NIED liability. The concern over too much litigation and liability is aggregate in nature. As such, it contains no distributive principle. Thus, it is not surprising that what we see by the end of the twentieth century is not a realization of the fault principle in the area of emotional distress, but the realization that the courts are obligated out of practical considerations to restrict its operation, even though the courts cannot justify all or even most of the lines that they are drawing. Instead, all they can do is acknowledge the need for some line to be drawn while being candid, as was the California Supreme Court in *Thing v. LaChusa*, that any such line will, in essence, be picked out of a hat.¹²⁸

¹²⁷ Prosser, *supra* note 91, § 34, at 212 (“[T]he only valid objection against recovery . . . is the danger of vexatious suits and fictitious claims, which has loomed very large in the opinions as an obstacle.”).

¹²⁸ 771 P.2d at 828 (acknowledging the necessity of drawing arbitrary lines limiting liability). The Supreme Court’s decision to review the *Ayers* case, 122 S.Ct. 1434 (2002) (mem.)—an unreported opinion that offered a fairly straightforward interpretation of FEOLA over which there is no circuit split—was perhaps driven by a sense that some device—whether a particularly narrow definition of “predicate injury,” or a

C. Against NIED

The standard account of the law pertaining to negligent causation of emotional distress is essentially an admission of failure, or at least an expression of substantial humility. The law in this area can do no better than announce restrictions on liability whose only justification is: "We can't do any better than this." Such a state of affairs at least warrants an attempt at reconceptualization. Has something gone unrealized that, once realized, might point to a more satisfactory analysis of these issues? We think so.

The problem with the standard view stems from its reliance on the equivalence thesis. This is *not* to say that the equivalence thesis is false. Rather, the equivalence thesis is not the appropriate benchmark for assessing the coherence of negligence jurisprudence (and tort law more generally). The equivalence thesis becomes the benchmark only if negligence law is conceived in a particular way, namely as a device for deterring and compensating victims of harmful conduct. It is that deeper premise, when conjoined with the thesis treating all species of harm as equivalent, which gives rise to the view that, *prima facie*, negligence law ought not to draw doctrinal distinctions between physical and emotional harms. Likewise, it is that deeper premise, which, by suggesting that negligence liability for emotional distress ought in *principle* to be wide open, requires courts and commentators to scramble to implement *ad hoc practical* constraints on such claims.

If one takes a different perspective on negligence, however, then one can grant the truth of the equivalence thesis, but still insist, as negligence law traditionally has, on distinguishing among species of harm. Part I of this Article suggests that an adequate explanation of the injury requirement necessitates a shift in the focus of analysis from the defendant to the plaintiff. The injury requirement is not well framed in terms of the question: "Did the defendant do something such that he is deserving of sanction or such that it would enhance social welfare to sanction him?" Rather, the right question to ask is: "Has the plaintiff been wronged such that she is

heightened proof requirement as to plaintiff having experienced fear—is necessary to ward off a flood of fear-of-injury claims in the asbestos context. It remains to be seen whether the Justices will decide to impose such a filter, and whether they can justify it on anything other than *ad hoc* policy grounds.

now entitled to seek recourse?" Now the standard analysis has again focused on the wrong party, although this time the problem manifests itself as an exclusive focus on the plaintiff, not the defendant. To be sure, a plaintiff must establish an injury to have standing to sue in negligence. But injury, of course, is not of itself sufficient to sustain a negligence cause of action. The plaintiff must further demonstrate that she suffered that injury because the defendant breached a duty to take care not to cause the type of injury of which the plaintiff complains. As a law of civil recourse, negligence law empowers a plaintiff to complain in a court that she has been injured as a result of the defendant failing to heed a duty to be vigilant of her well-being. The critical question then becomes: "Of what sort of harm was the defendant under a duty to be vigilant?" With this shift in perspective, it becomes quite intelligible to differentiate among species of harms. To deny that emotional distress is generally actionable does not require us to deny that it is a harm, nor to denigrate it as a less serious or important form of harm than physical injury. It is instead to assert that actors are not ordinarily under a legal duty to be vigilant of others' emotional well-being.

We have seen, as a matter of black-letter law, that this latter proposition about duty seems quite well-established. Whereas decisions such as *Heaven v. Pender*¹²⁹ and *MacPherson v. Buick Motor Co.*¹³⁰ long ago announced a general duty to take care not to cause physical harm and illness, no comparable decisions have taken hold with respect to emotional distress.¹³¹ This is just another way of saying that there is no such thing as a general principle that negligently inflicted emotional distress is an actionable tort. Instead, there is a patchwork of limited-duty rules.

As with the black-letter requirement of injury, the reader will no doubt want some explanation for the absence of a general duty to avoid causing emotional distress. Why would the law acknowledge a general duty to be vigilant of others' physical well-being but not an equivalent duty to be vigilant of others' emotional well-being? We have briefly canvassed some of the reasons that support this

¹²⁹ 11 Q.B.D. 503 (Eng. C.A. 1883).

¹³⁰ 111 N.E. 1050 (N.Y. 1916).

¹³¹ Professor Dobbs notes that two state supreme courts, Montana and Tennessee, have haltingly flirted with the proposition. See Dobbs, *supra* note 19, § 312, at 850–51.

distinction elsewhere,¹³² as Professor Geistfeld has done, from a somewhat different perspective, in his contribution to this Symposium.¹³³ These include the idea that negligence law must set duties in a manner that permits actors to set priorities about the sorts of injuries they must take care to avoid, which are not overly demanding. We elaborate and expand on these explanations below.

Before doing so, however, one more bit of brush clearing is in order. This concerns the traditional distinction between predicate injury and parasitic damages. As we have noted, the distinction is a very real one in the law.¹³⁴ Lost profits are typically recovered as parasitic damages in a personal injury action, but negligent causing of pure economic harm is generally not actionable. An unforeseeable extent of damages is actionable as parasitic damages consequent upon a foreseeable physical injury, but where the predicate injury is itself unforeseeable, there might be no action at all.

One of the most familiar of these distinctions is drawn in the area of emotional harm. Pain and suffering, trauma, and fear are generally proper subjects of damages in a personal injury action based on physical injury. On the other hand, as we have seen, negligently causing pure emotional injury, outside of a few pockets, is generally nonactionable.¹³⁵ Strikingly, courts sometimes apply a version of this type of rule in the fear of disease scenario. An individual who has been diagnosed with benign mesothelioma and consequently has a fear of cancer may recover for the actual physical injury of mesothelioma and the emotional injury of fear of cancer. On the other hand, a person with the same exposure and no mesothelioma often will be unable to recover for the fear of cancer.¹³⁶

All of this is quite unintelligible on the standard account. If there is liability for fear of cancer, it should be because fear of cancer is foreseeable and is the sort of injury for which recovery is provided. But, if this is true, then there should be liability for fear of cancer regardless of whether it accompanies a physical injury for which

¹³² Goldberg & Zipursky, *supra* note 4, at 1833 (discussing prioritization of obligations).

¹³³ Geistfeld, *supra* note 69, at 1929-34 (discussing prioritization from an economic perspective).

¹³⁴ See *supra* text accompanying notes 87-92.

¹³⁵ See *supra* text accompanying notes 87-103 (describing the actual exposure rule).

¹³⁶ See *supra* text accompanying notes 87 & 104.

there is liability. The only explanation for the predicate/parasitic distinction must, again, be that it functions as a crude limit on excessive or fraudulent litigation.

It is a mistake to suppose that, where there is the same act and the same consequence, there must be the same possibility of recovery. The problem is that there are many ways of defining an act. One way to define an act is in part by reference to its consequences. Whether there has been an actionable breach of duty in negligence law depends upon whether the defendant has in fact performed an action that injured the plaintiff. Not every such action—not even every carelessly performed act of injuring the plaintiff—is actionable. Rather, the injury caused by the defendant must be an injury that the defendant has a duty to be careful not to cause.¹³⁷ The general rule in tort law is that one has a duty to take care not to cause physical injury to others, and that one does not have a duty to take care to prevent emotional injuries or pure economic injuries to others. If the wrong species of injury is realized, then there is a mismatch between the duty pleaded and the injury proved: The injury complained of is not one that the defendant was under a duty to take care not to cause.

But if this is so, why do courts permit damages for emotional harm even as parasitic damages? The answer is the traditional one: The measure of damages is an entirely different question from what sorts of duties one must observe to avoid injuring others. Rather, the measure of damages is typically made by ascertaining what would fairly compensate a plaintiff who has made out a breach of duty causing the right kind of injury. This is a question about remedies, not about whether there is an actionable tort. Conversely, where there is no predicate tort of negligently causing physical injury, there should be no liability for mental injury.

An implication of this critique is that the ancient phrase *damnum absque injuria* (harm, but not cognizable harm) may help illuminate the area of emotional harm. For it is plain that, where

¹³⁷ Traditional formulations of the negligence tort capture this limitation by defining it, in part, as the “breach of a duty of reasonable care owed to persons such as the plaintiff.” This is opposed to the looser, modern formulation, “failure to act as a reasonable person.” See Goldberg & Zipursky, *The Place of Duty*, *supra* note 6, at 684–87 (noting that abandonment of “duty” and “breach” causes misdescriptions of the negligence tort).

there has been emotional distress due to fear of a disease or fear of injury, there has been some form of harm to the plaintiff. That is why something more than compensation for physical damages is properly required to make the plaintiff whole. But it does not follow from the existence of emotional distress caused by wrongful conduct that there has been a wrong. In our system, it is required that the emotional harm rise to the level of an emotional "injury" if the negligence causing it is to count as a wrong. The concept of an "injury," used in this way, is normative in a way that "harm" or "damages" is not. It requires that the loss that the plaintiff has suffered is a loss that she has some kind of entitlement to be free from, a part of her in which she is entitled to be secure. Correlatively, the defendant must have been under a duty to take reasonable care not to cause this type of injury.

D. Rethinking NIED: Agency and Responsibility

Courts have struggled in their attempts to analyze fear-of-injury cases within the NIED framework. This is not surprising, we have suggested, because the NIED framework is itself built on a set of misconceptions. The law of negligence does not in fact recognize a general obligation to be vigilant of others' emotional well-being that needs to be cabined-in by various ad hoc restrictions. None of this is meant to suggest, however, that the courts have been mistaken in permitting certain plaintiffs who fear future injury from recovering from defendants who have negligently created that fear. Rather, it suggests that we need a better way of accounting for these claims, and for determining when departures from the background rule are acceptable.

In the following Sections, we develop an alternative analytic framework for cases of negligence causing fear of future illness. Because the argument is somewhat intricate, it is worthwhile to provide a brief road map. The previous Section criticized the standard view that the default rule against recovery for emotional harm, as well as its exceptions, are explicable in terms of various concerns that fall under the general heading of administrability. Section III.D.1 criticizes the most common and influential of these arguments: the fraud argument. That argument maintains that the default rule is prompted by a concern that the subjectivity of emotional harm—the fact that it is an internal mental state not directly

observable to others—renders negligence law prone to abuse by plaintiffs who will feign psychic injuries to obtain compensation. The fraud argument is unconvincing as stated. Its focus on the subjectivity of emotional distress, however, points to a different and more plausible concern that does help explain negligence law's suspicion of emotional harm claims: Emotional harm is often, though not always, in the control of, and therefore the responsibility of, the victim. We call this the "agency concern." When properly fleshed out, the agency concern not only yields an understanding of why negligence law is reluctant to acknowledge emotional harm claims, it also explains why courts have drawn exceptions to the default rule of no liability in certain clusters of cases.¹³⁸

Section III.D.3 brings together Part II's discussion of heightened risk with the discussion of emotional harm in order to reveal the nature of the dilemma that courts have faced in the future injury cases. Recognition of heightened risk of future illness as an injury in itself is problematic because it seems to render cognizable a claim based on something which, if it is a harm at all, seems to be a premature or unripened version of the "real" harm at issue (the harm of a serious illness). Yet, recognition for claims of emotional harm associated with heightened risk of illness seems to invite claims for a type of harm that is, in a normative sense, the victim's doing or the victim's problem. The Section then advances a philosophical framework that permits us to navigate the horns of this dilemma and thereby explain why some claims for negligence causing emotional distress resulting from exposure to a heightened risk of disease ought to be cognizable. Specifically, it harnesses the concept of "secondary legal qualities" to explain how certain interferences with one's life can be objective in the specific sense that they have the characteristic of being the sort of interference that tends to cause reasonable persons who are exposed to them to suffer emotional distress. Section III.E pulls these various argumentative strands together by arguing that a defendant who carelessly generates a set of circumstances that interfere with the plaintiff's interest in being free from the pall associated with being at substantial risk of a serious illness may be held responsible for the emotional distress experienced by the plaintiff because of those circumstances.

¹³⁸ See discussion *infra* Section III.D.2.

The pall of being under a threat of serious illness is an objective state of affairs that can be identified by judges. It also attests that the victim's emotional distress is not fairly described as being her problem or of her own making. In short, we argue that "future injury" claims can be cast as claims for negligently causing the victim to endure a significant threat of suffering a serious illness.

1. Fraud and Subjectivity

In 1941, Dean Prosser maintained that "the only valid objection against recovery for mental injury is the danger of vexatious suits and fictitious claims."¹³⁹ This assessment, still widely shared, rests on the idea that emotional harm is fundamentally subjective in a way that bodily injury is not. So put, the fraud concern is remarkably unpersuasive, as we demonstrate below. Why, then, have so many legal scholars been attracted to it as an account of the limited nature of obligations to avoid causing emotional distress? The fraud (or feigning) argument has persisted because it captures, clumsily, a different and deeper concern with liability for emotional harm. We label this deeper concern the "agency concern," and develop it in the remainder of this Section. Understanding the agency concern is essential to realizing what is going on in cases raising issues of NIED.

The fraud objection to general recovery for negligent infliction of emotional distress is unpersuasive for several reasons. First and foremost, most of the reported cases involve factual scenarios where there is every reason to believe that the plaintiff did or would have suffered some emotional harm.¹⁴⁰ Thus, for example, a woman frightened by horses galloping right up to her face probably would suffer at least some negative effect, for example, fear. Any person who is at heightened risk of cancer probably will suffer some distress from knowing this fact; likewise for persons facing a substantially heightened risk of HIV infection. In cases where a child, spouse, friend, or even a stranger is injured or killed in (or sometimes out of) the presence of the plaintiff, there is little doubt that some emotional harm exists. In these scenarios, there are, of course, issues whose veracity merits scrutiny. One issue concerns

¹³⁹ Prosser, *supra* note 91, § 34, at 212.

¹⁴⁰ Dobbs, *supra* note 19, § 302, at 822-23.

whether events in fact unfolded in the manner alleged by the plaintiff. This issue, of course, is not unique to claims for emotional distress. Another issue is whether the extent or degree of emotional harm alleged by the plaintiff actually exists. This latter question, however, is also present in cases allowing recovery for emotional harm as parasitic damages, and the common law has a general rule permitting such awards.

The asserted risk of fraud supposedly emerges from the fact that emotional harm is subjective—that is, in the plaintiff's mind. But it is far too late in the day, either in torts or in the law more broadly, to eliminate a legal concept or remedy because it raises an issue of subjective mental states. Every case alleging an intentional tort, seeking punitive damages, or pursuing a claim of defamation on behalf of a public figure requires the fact finder to make some determination as to a party's state of mind, yet that does not present an obstacle in these areas. Moreover, as to particular emotional harms for particular individuals, it is today possible to utilize specialists in psychiatry and psychology who are generally capable of ascertaining, at least to some degree, the extent of emotional damage.

Finally, if the fraud objection were serious, one would expect changes in the law to reflect this. They do not. Consider, for example, the weakening of the physical injury rule into a physical impact rule. It is probably as easy to claim a slight physical impact—a touching—as an emotional harm. And it is far easier to fabricate an impact than an actual injury. Consider also that many jurisdictions that permit recovery for emotional harm do so only when it is severe. If plaintiffs can feign psychic injury to a degree that renders such claims unadjudicable, is there any reason to suppose that they are incapable of disingenuously converting a minor psychic injury into a severe one?

We do not raise these challenges to the fraud objection out of a desire to open the way to a wide array of emotional distress claims. Rather, we suggest that if the fraud objection is as weak as it appears to be, yet continues to grip courts and commentators, it must be because there is some other idea at work, related to the fraud objection, that actually does carry some weight.

For the purpose of elucidating this idea, it will be helpful to draw an analogy from a seemingly distant realm outside the adjudication

of legal disputes. Suppose a child returns from school complaining of hurt feelings or emotional injury because of schoolyard taunts. Her parent(s) might respond in a number of ways. One possibility is to focus on the child's feelings, to explore their dimensions, and to offer solace and intangible or even tangible acts designed to foster emotional healing (for example, a hug or a trip to Ben & Jerry's).

Often, however, there is a balance to be struck, for it might well be a mistake to foster an undue preoccupation in the child with her hurt feelings. The problem is that children sometimes overreact to negative experiences like the one described. Sometimes they make a mountain out of a molehill. In certain contexts, therefore, it may be appropriate for parents to downplay the hurt feelings, to change the subject, to foster an attitude of shrugging it off. The parent who utters it, more realistically, conveys in other than words, clichés like "Sticks and stones may break my bones, but names will never hurt me" or "Don't make a mountain out of a molehill" is taking this tack.

To tell a child that she is making a mountain out of a molehill might sound like an accusation of misrepresentation, and sometimes it might be used to convey that message. But its literal meaning assumes that in fact there is something to complain about (the molehill). Thus, when the expression is invoked, it typically does not carry the suggestion that the child is making things up. Rather, it is meant to inculcate or reinforce a norm of how to respond, emotionally, to an adverse event by being resilient or thick-skinned. A parent who takes this approach aims to provide a sort of emotional education or nurture a kind of spirit, not to accuse his child of fraud. A culture, institution, or school that does not tolerate children crying beyond a certain age is not involved in disbelief of an emotional response; it is simply involved in setting norms.

The directive "Don't make a mountain out of a molehill" not only aims to guide the child toward a response of minimization. It also asserts, at a deeper level, that the emotional response is something that is, to some degree, in the child's own control. In an unusually literal way, it presupposes that the child is in part responsible for his emotional response.

Now let us return from the home and the schoolyard to the courtroom. We believe that the much-vaunted fraud objection is in

fact a loose way of expressing the concerns that particular plaintiffs in emotional distress cases are making mountains out of molehills, and that permitting a cause of action for pure emotional harm will invite or encourage citizens to make mountains out of molehills.¹⁴¹ As the foregoing analogy suggests, it is a mistake to interpret the core of this idea as an accusation of, or concern about, fakery or feigning. Rather, it refers, albeit obliquely, to two other things we have mentioned. First, it provides a legal version of norm-setting and norm-articulation, where the norm is one of fostering the thick-skinned response of getting on with life, rather than dwelling on one's upset. Second, it is a recognition that, to some degree, the plaintiff herself bears responsibility for the depth of emotional harm she feels. This is not to say that, as a matter of fact, all persons have a choice in each instance as to how great an emotional effect some event or accident will have. Rather, it is a normative ideal of holding plaintiffs responsible for their personal responses.

In a nutshell, then, the idea is that a plaintiff's emotional response to a state of affairs is a matter for which she herself can sometimes, indeed often, be held responsible. Emotional distress is not simply an uncontrolled or uncontrollable reflex, as is the bodily response to a blow or a toxin. It is a response mediated by the mind of the plaintiff. Despite its obvious vices, the old rule of proximate cause actually captured this idea. It went astray because it simply assumed that the existence of this mediation was *always* sufficient to cut off responsibility on the part of the defendant.¹⁴² In this assumption, it took on the cast of a harsh "blame the victim" mindset, characteristic of an era captivated by rugged individualism. We will suggest, however, that a more subtle version of the responsibility idea does not translate into a blame the victim rule for emotional distress claims. Rather, it will help us to identify those instances in which plaintiffs seeking recovery for emotional distress ought to recover.

Recognition of the notion that a plaintiff's agency plays an important role in the generation of emotional distress has several implications for the doctrinal structure of negligence law. First, it is

¹⁴¹ Thanks to Professor Margo Schlanger for encouraging this line of inquiry in the course of discussions at the conference.

¹⁴² See *supra* text accompanying notes 90–91 (describing the old proximate cause doctrine).

not clear that severe emotional harm is always something for which the defendant whose conduct precipitates the harm is responsible. If, in fact, the escalation into emotional distress is something the plaintiff is entirely responsible for, then the defendant is off the hook. Second, and more saliently for the vast run of cases, it suggests a need to identify conditions or circumstances that will help determine when a plaintiff can be said to be not responsible for his experience of distress. These points, taken together, may further suggest that fear-of-future-injury claimants are not, after all, pursuing claims for "pure" emotional distress. Rather, they are complaining about the presence, in plaintiff's life, of a state of affairs that precipitates emotional harm, and does so in a manner that does not so involve the plaintiff's responsibility that it calls into doubt the cogency of assigning responsibility to defendant.

If these observations are correct, then the absence of a general duty to avoid causing emotional distress links up to certain other features of negligence law, including comparative fault, assumption of risk, and the duty to mitigate damages. Each of these features denies or reduces plaintiff's recovery on the ground that plaintiff is wholly or partially responsible for his injury. Comparative responsibility and failure to mitigate generally involve a notion of plaintiff fault. Assumption of risk asserts that the plaintiff is responsible for certain injuries regardless of fault, because she knowingly and freely chose to take on certain risks of others behaving carelessly. Just to be clear—we do not mean to suggest that emotional overreaction is the fault of the plaintiff, or that plaintiffs typically "assume the risk" of being distressed. We do not believe that the idea of responsibility for distress is, or should be incorporated into, the doctrines of comparative fault or assumption of risk. At this point, we are merely noting that the agency concern bears a family resemblance to those other ideas.

The connection between the agency concern and the default rule against recognizing a duty to take care not to cause emotional harm can be illuminated by considering the traditional rule barring licensees from suing landowners in negligence for physical injuries caused by unsafe conditions other than known traps. Both emotional distress claims and licensee claims are frequently categorized as "limited duty" cases. Both depend upon the idea that the defendant in these classes of cases should not be held responsible for this

aspect of the plaintiff's well-being. This is not—as in duty-to-rescue cases—because of the absence of misfeasance on the part of the defendant. Nor is it because the assignment of liability in these instances either would leave defendants shouldering responsibilities much more demanding than those customarily recognized of persons in similar situations or would expose them to conflicting obligations. Instead, the courts' circumscription of these duties is itself an allocation of responsibility for certain aspects of plaintiffs' well-being to the plaintiffs themselves. Although this analysis in some ways invokes ideas similar to that of assumption of risk, there is no suggestion that these plaintiffs knew of, or voluntarily subjected themselves to, the risk of the defendant acting carelessly toward them. Instead, when the courts in these areas deny a cause of action, it is based on their conclusion that the obligation to take care to avoid injury to the plaintiff ought to be assigned to the plaintiff. Because this allocation concerns responsibility in a sense that pertains to the scope of obligation, it is rightly expressed in terms of "no duty" or "limited duty" rather than assumption of risk. Thus, the traditional rule for licensees reflects a judgment, right or wrong, that it shall be the licensee's obligation to see to it that she is not injured by unknown, unsafe conditions when she is a social guest on another's land. Similarly, the default rule against recovery for emotional harm reflects a judgment that the maintenance of one's emotional well-being in the face of adversity is something for which a plaintiff ordinarily must take responsibility.

The idea of responsibility for one's own emotional responses also links to two other aspects of negligence law. One of them is the irrelevance of diligence and good faith to the inquiry into the defendant's breach and the plaintiff's comparative fault. As Justice Holmes famously demonstrated, the common law of negligence holds each of us to a standard of conduct outside of our own particular strengths and weaknesses.¹⁴³ Similarly, the general rule against recovery for pure emotional harm borrows from this objective standard. In denying some emotional distress claims, the law does not suppose that plaintiffs are fabricating their distress. Rather, it holds them to an external standard that, to some extent, ignores their particular vulnerabilities.

¹⁴³ Holmes, *supra* note 2, at 108.

Finally, and perhaps most importantly, negligence law's evident wariness of assigning liability for pure emotional harm connects with the very idea of civil recourse. As Professor Weinrib has untiringly argued, tort law conceives of the plaintiff and defendant as having been enmeshed in a certain kind of episode in which the defendant is the "doer" and the plaintiff the "sufferer."¹⁴⁴ In this episode, the defendant is the agent of harm and the plaintiff the victim of harm. The defendant is an actor who acts upon the plaintiff and inflicts an injury upon her.¹⁴⁵ This conception of tort as a paradigm involving the defendant *actively* wronging the plaintiff is stylized in certain ways and we have criticized it elsewhere.¹⁴⁶ But it captures the core truth that our system does *not* treat plaintiff and defendant as reciprocal causes—this idea is precisely where the radicalism of Professor Coase and Judge Calabresi lies.¹⁴⁷ Instead, tort conceives of the plaintiff with a meritorious claim as a victim rather than an agent. The notion of the plaintiff as one who is *acted upon* plays a central role in the system of civil recourse, for it is being acted upon that triggers the right to redress one's injuries. As a wide array of reputational, dignitary, and pecuniary "injuries" in tort indicate, there is nothing essentially physical or bodily about this conception of being acted upon, of being the victim of a wrong. Nevertheless, what is crucial—at least historically under the common law—is that the injury be inflicted from without, rather than self-inflicted. The plaintiff's injury cannot be her responsibility, for she is using the legal system to obtain recourse for something done

¹⁴⁴ Weinrib, *supra* note 56, at 81.

¹⁴⁵ *Id.* at 63–66. Weinrib ties the conception of what happens when a tort occurs to an ideal of "corrective justice," in which tort law is viewed as reversing or undoing the wrong visited on the victim by requiring the injurer to return what he has taken from the plaintiff. *Id.* at 64–65. We have here and elsewhere rejected this characterization of how tort law responds, suggesting that it does not so much force the defendant to give something back to plaintiff, as it permits plaintiff to seek recourse (through the law) against the defendant. Nonetheless, Professor Weinrib's characterization of individual torts as essentially involving an injurer visiting a wrong upon a victim is quite sound.

¹⁴⁶ See, e.g., Zipursky, *supra* note 49 (distinguishing civil recourse from Weinribian corrective justice theory).

¹⁴⁷ This, of course, is a point that Professor Epstein long ago conveyed in a somewhat different context. See Richard A. Epstein, *A Theory of Strict Liability*, 2 J. Legal Stud. 151, 160–61 (1973) (criticizing Coase's conception of causation as applied to tort law).

to her by someone else. Characterizing the plaintiff's emotional response as her responsibility undercuts the idea that defendant inflicted it.

2. Revisiting Exceptions to the Default Rule of No Duty

The reader should again bear in mind that the account here is not intended to justify a blanket rule against recovery for pure emotional harm. Rather, it is intended to explain the reluctance of the courts to recognize a broad duty to take care to avoid causing emotional distress. The nub of the explanation is that we do not treat the plaintiff's emotional harm as an injury that exists independent of the plaintiff's own response or independent of the legal system's choice to treat it as such. Instead, the rule is itself a norm that holds persons responsible and refuses to deem emotional injury as a fact impervious to the plaintiff's own will.

The notion of bearing (at least partial) responsibility for one's own emotional state helps illuminate several aspects of negligence law. First, it helps make sense of both the predicate physical injury rule and why the judicial attempt to water-down the predicate-injury rule into an impact rule is fundamentally wrongheaded. The predicate-injury rule reflects a generalization that, when one suffers distress in connection with a traumatic physical harm or illness, one does not make a mountain out of a molehill. Taken in reference to substantial physical wounds and illnesses, the rule has some plausibility. That plausibility diminishes, however, when the rule tends toward overinclusiveness, which is exactly the consequence when one substitutes impact for injury. The "least touching" formulation, long applied to intentional wrongdoings such as many batteries, has no place in negligence precisely because it fails to take seriously the fact that a victim's emotional distress over slight touchings is, absent other circumstances, her own responsibility.

The notion of responsibility for emotional states also explains other contours of tort. The tort of assault is thought to provide the oldest example of a recognized claim for "emotional distress."¹⁴⁸ Assault is actionable when the injurer causes an apprehension of imminent harmful or offensive touching in the victim. When a bullet or fist whizzes past someone's head and he or she feels fright,

¹⁴⁸ Prosser, *supra* note 91, § 10, at 48 (noting old actions for assault).

that is an emotional response, but it is quite different from the example of the schoolyard taunt. The response is visceral, immediate, and unthinking. In this context, it makes little sense to hold the plaintiff responsible for the response and makes much more sense to think of the plaintiff as a victim who exercised little or no agency. Relatedly, the idea of responsibility yields insight into the notorious "zone of danger" rule. A person who is in a car and experiences a car crash, or is otherwise part of an episode of grievous peril, will often respond to that episode viscerally.¹⁴⁹ Indeed, it is quite unreasonable to expect anything other than a visceral response to an episode of extreme imperilment and serious injury to another. This is not to say that the zone of danger rule is the appropriate standard for claims of negligence causing emotional distress, but only that its role as a longstanding and conspicuous exception to the general no-duty rule is interestingly explained by the foregoing account.

Third, many contemporary cases contemplate relaxing the bar on recovery for pure emotional harm only where the harm is "severe."¹⁵⁰ Indeed, there is now a principle or catchphrase that is commonly trotted out to capture this notion of severity. Consider the following comment on "severe emotional distress" from the *Restatement (Second) of Torts*:

Emotional distress passes under various names, such as mental suffering, mental anguish, mental or nervous shock, or the like. It includes all highly unpleasant mental reactions, such as fright, horror, grief, shame, humiliation, embarrassment, anger, chagrin, disappointment, worry, and nausea. It is only where it is extreme that the liability arises. Complete emotional tranquility is seldom attainable in this world, and some degree of transient and trivial emotional distress is a part of the price of living among people. The law intervenes only where the distress inflicted *is so severe that no reasonable man could be expected to endure it*.¹⁵¹

¹⁴⁹ See, e.g., *Robb v. Pa. R.R. Co.*, 210 A.2d 709, 709-10 (Del. 1965) (allowing recovery for NIED when the plaintiff narrowly avoided being run over by a train).

¹⁵⁰ See, e.g., *Kosmacek v. Farm Serv. Co-op of Persia*, 485 N.W.2d 99, 104 (Iowa Ct. App. 1992).

¹⁵¹ *Restatement (Second) of Torts* § 46 cmt. j (1965) (emphasis added).

While this comment pertains to intentional infliction of emotional distress, its language has found its way into some negligence cases in the fear of disease area.¹⁵²

There is something very peculiar, on first blush, about the language of this comment, for it is not clear what it means to be unable to "endure" emotional distress. Does this language mean that only a person who has "fallen apart" may recover? If not, what does it mean? This puzzle only arises because of a slip in the courts' usage of the "no one could be expected to endure" language. It is not the emotional harm itself that plaintiffs cannot be expected to endure. After all, plaintiffs recover even if they do endure tremendous fear, grief, upset or anguish, at least if "endure" means survival at a functioning level. The problem is in the composition of the language of the comment. It is not the emotional *distress* that must be "beyond what a reasonable person could be expected to endure," it is the *stress* that must be beyond what that person could endure. If this stress exceeds a certain threshold, then a reasonable person cannot be expected to endure it. There are situations in which a reasonable person cannot be expected to keep a stiff upper lip; she cannot be expected to avoid responding by tumbling into severe emotional distress, that is, situations in which the mountain is a mountain, not a molehill.

We can illustrate this last point by taking a particularly gripping example, one which provided the facts that led the New Jersey Supreme Court to modify the zone of danger rule in *Portee v. Jaffee*.¹⁵³ The plaintiff was a mother who stood outside an elevator shaft for four hours as her five-year-old son moved up and down the shaft, caught between the elevator doors and the shaft, screaming.¹⁵⁴ Hours after the elevator was finally stopped, the child died.¹⁵⁵ The mother was severely traumatized by this event and sought recovery for emotional harm.¹⁵⁶ The defendant building owner argued that, since the mother was not in the zone of danger, she should not be

¹⁵² E.g., *Doe v. Northwestern Univ.*, 682 N.E.2d 145, 150 (Ill. App. Ct. 1997), *aff'd* on other grounds, *Majca v. Beekil*, 701 N.E.2d 1084 (Ill. 1998); *Kosmacek*, 485 N.W.2d at 104.

¹⁵³ 417 A.2d 521 (N.J. 1980).

¹⁵⁴ *Id.* at 522.

¹⁵⁵ *Id.* at 523.

¹⁵⁶ *Id.*

able to recover.¹⁵⁷ The New Jersey Supreme Court rejected the argument and adopted a rule permitting bystander recovery.¹⁵⁸

Portee presents a compelling situation in which no parent could reasonably be expected to endure without suffering severe emotional distress. This is not to say that her level of subjective distress was more than one could be expected to endure. It is rather that the experience of helplessly witnessing for several hours the accidental maiming and killing of one's child is "more than a reasonable person could be expected to endure." There is no place, in a case like this, for a court to send the message, "get over it" or "keep a stiff upper lip." This is not because the emotional harm is particularly severe as compared to others' emotional distress, but instead because of the severity of the circumstances giving rise to it.

The explanation for why certain cases present actionable wrongs is symmetrical with the explanation of the general no-cause-of-action rule. In the normal case, the norm is to treat the emotional harm as, in part, the responsibility of the plaintiff; it does not permit her to treat the injury as a fact entirely beyond her control; society literally *expects* her (in a sense more normative than predictive) to endure her circumstances rather than to plummet into great emotional distress. It is quite natural then that in a case where a reasonable person could not be expected to endure the circumstances, the emotional harm is an injury for which the plaintiff may be treated as a victim rather than as an agent.

3. *Evading the Dilemma of Fear-of-Future-Disease Claims: Secondary Legal Qualities*

In cases similar to *Portee*, the analytic route around the agency concern is to point to features of the situation that so overbear the plaintiff's resilience that one cannot regard the plaintiff's being emotionally shattered as his responsibility. However, we have *not* highlighted these cases in aid of establishing a new threshold of "unendurability" over which all plaintiffs must pass in order to recover in negligence for emotional distress. Rather, we have highlighted the issue of "endurability" to help capture a central concern

¹⁵⁷ *Id.*

¹⁵⁸ *Id.* at 528.

that lurks behind the courts' disinclination to recognize negligent infliction of emotional distress claims, as well their willingness to abandon that disinclination in certain circumstances.

In fact, as we noted at the outset of this Part, negligence law not only suppresses its general hostility to emotional distress claims in those situations in which it is not fair to ask of the plaintiff that she exercise agency with regard to her emotional response (as in *Portee*), it also permits claims when the plaintiff's emotional response is not really the primary harm about which she is complaining. A classic example of this is a cause of action for libel, where the libeled party has not suffered any economic loss. Typically, one who has been falsely branded a child molester, for example, will suffer emotional harm without physical harm, and the emotional harm is part of why she is suing. Yet there is still a predicate harm in this case—the intangible predicate harm of reputational damage—that provides the foundation for the cause of action.

These reflections on the role of intangible predicate harms return us to the earlier analysis, in Part I, of claims for heightened risk of injury. In particular, it is tempting to argue that plaintiffs who are placed at heightened risk of physical injury have suffered a kind of intangible predicate harm, like harm to reputation, that could support a claim sounding in negligence for emotional distress. The problem with this argument is that, as stated, it amounts to bootstrapping. A claim for “pure” emotional distress is problematic because of the agency concern. A claim for heightened risk—even if that risk is intelligible as a harm—does not invoke the sort of harm that defendants have a duty to take care to avoid causing. How can one make the case for recognizing a cause of action for fear of future injury by combining these two different kinds of problematic claims? To pose a more constructive version of the same question: Is there a way that these two conditions—circumstances posing a heightened risk of physical injury to the plaintiff and the existence in the plaintiff of a subjective emotional response—might combine to generate a basis for a claim in negligence for fear of future disease? We will argue that there is a way. In order to flesh out this idea, however, we must pause to sketch a concept developed in another context.

In an earlier article on self-defense, one of us argued that certain legal categorizations are best understood as identifying qualities

that are akin in their structure to what John Locke called "secondary qualities."¹⁵⁹ These are qualities whose possession consists in their being such as to be responded to by people in certain ways. The classic example, for Locke, was the secondary quality of a color—for example, the color blue.¹⁶⁰ The color blue is an attribute whose existence consists of looking blue to a normal person in ordinary lighting. Of course, blue things also have certain underlying attributes that make them look blue in ordinary lighting, to the normal perceiver. But their blueness, even if scientifically understandable at some different level, consists in their being such as to look blue.

The self-defense article suggested that certain qualities of actions or states of affairs are secondary in the Lockean sense that their nature consists of how a normal (that is "reasonable") person would respond to the action possessing the quality.¹⁶¹ In the case of self-defense in the criminal law, conduct by the aggressor that threatens imminent death or serious bodily harm triggers a right to take potentially lethal defensive actions. Whether the aggressor's conduct possesses this quality of threateningness depends on whether the aggressor deliberately engaged in conduct that a reasonable person would regard as presenting a high risk of imminent death or serious bodily harm. Thus, for example, an aggressor who points a real gun at the head of another and begins to pull the trigger has engaged in threatening conduct, even if the gun is not loaded. By contrast, an aggressor who points a banana at the head of another has not engaged in threatening conduct, even if the intended victim is psychotic and, through no fault of his own, believes the banana to be a gun. The legal claim was that, while a self-defense *excuse* may exist in the second case, a *justification* does not. The right to self-defense, under our law, is predicated on the existence of objectively threatening conduct.

The theoretical advantage of casting the quality of threateningness as a secondary legal quality is that it explains how the right to self-defense can turn on something *objective*—not solely in the mind of the putative self-defender, but in the conduct of the ag-

¹⁵⁹ See Benjamin C. Zipursky, Self-Defense, Domination, and the Social Contract, 57 U. Pitt. L. Rev. 579, 600 (1996).

¹⁶⁰ *Id.*

¹⁶¹ See *id.* at 601–03.

gressor—but nevertheless recognize that the relevant legal classifications are ineliminably anthropomorphic and contextualized at several levels, and more particularly, that law must reference the way normal individuals perceive or respond to actions, rather than to underlying qualities of the conduct or events that can be specified independently of the tendency of normal individuals to respond. Locke negotiated this metaphysical dilemma by defining secondary qualities. The law resolves a similar dilemma by retaining objectivity and qualities outside the mind while accepting an essential dependency on response-based categorizations, by utilizing “secondary legal qualities.”

Interestingly, the tort of assault also seems to employ a secondary legal quality. An assault is an intentional act that produces an apprehension of imminent harmful or offensive touching. The quality of being the sort of act that produces an apprehension of an imminent harmful touching in a reasonable person is the quality that makes the act an assault. Whether the act has this quality or lacks it is an objective matter. This is not to say that people can never disagree as to whether that quality is present or absent in a given case. Rather, it is to say that the quality is not something in the mind of the victim, but is rather an aspect of the act itself, albeit an aspect whose nature and existence is defined by a matching subjective response.

In a different way, the tort of public disclosure of a private fact also draws upon what we are calling a secondary legal quality.¹⁶² Whether something is a private fact is in significant part a secondary legal quality under the law: It is a fact, the disclosure of which would be highly offensive to a reasonable person. Again, the attribute in question is an attribute of the disclosed fact (for example, the fact that the plaintiff has genital herpes): it is the attribute of being the sort of information, the public disclosure of which would be highly offensive. And yet the essence of the attribute is specified in terms relating to the response—in this case the emotional response—of a reasonable person.

The broader lesson to be drawn from this discussion is that the harm in certain valid emotional harm cases is not best cast as damage to something *in the plaintiff*, to wit, her emotional well being.

¹⁶² See *id.* at 600–01.

The injury instead should be cast as interference with the plaintiff's interest in avoiding certain kinds of circumstances. Once one perceives this, the case law of emotional distress appears less odd and more aligned with other aspects of tort law. Consider, for example, the tort of false imprisonment. That tort involves intentionally creating a situation in which a plaintiff would reasonably take herself to be confined.¹⁶³ Freedom from confinement is, needless to say, a substantial liberty interest shared by all. Certain acts unjustifiably interfere with this interest and are therefore invasions of the plaintiff's right, as well as being "wrongs" under the tort law. Of course, the primary effect this tort has on an ordinary person is to cause them to experience fright, anger, embarrassment, frustration, and/or other negative feelings. But a cause of action for false imprisonment does not identify these as the injuries, nor does it identify the causing of them as the wrong. Instead, false imprisonment is a dramatic example of depriving someone of an important liberty, the liberty of movement. The essence of the injury is the deprivation of that liberty.

A similar account can be offered for other torts. Consider, for example, defamation. The essence of the tortious conduct is the tarnishing of the plaintiff's reputation, and the injury is an injury to reputation.¹⁶⁴ Emotional harm is a predictable consequence, and it will sometimes be the most obvious consequence and provide the most significant damages in a defamation action. It is the interference with the reputational interest, however, that constitutes the actionable injury.

Or take the example of nuisance. The tortious wrong is an interference with a proprietary interest, the interest in use and enjoyment of one's property.¹⁶⁵ Apart from injunctive relief, the classic remedy for nuisance has been damages equal to the diminution in value of the property caused by the nuisance. Nevertheless, for many decades damages for emotional harm have also been available, on the theory that the disruption of enjoyment of one's life caused by nuisances are actionable. The character of the remedy notwithstanding, the wrong of nuisance is not the infliction of emo-

¹⁶³ See Dobbs, *supra* note 19, § 37, at 67-68.

¹⁶⁴ See *id.* § 400, at 1117.

¹⁶⁵ *Id.* § 463, at 1321.

tional distress itself, but the interference with the plaintiff's property interest.

Privacy torts consist of the same dual nature. No doubt, what animated Warren and Brandeis' thinking, at least in part, was the significance of the harm to the individual's sensibility caused by invasions of privacy.¹⁶⁶ But the wrong of the tort is not itself the infliction of the harm. The tort instead involves interference with a particular kind of liberty interest: a freedom from intrusions of a certain sort, and as Dean Prosser later elaborated, an interest in freedom from disclosures and misappropriations.¹⁶⁷ The wrongs themselves are definable in terms of the defendant's objective acts that are realized in certain kinds of actual interference with the plaintiff's interests, even though that interference is typically accompanied by certain kinds of subjective harm as well.

Finally, it is worth noting that many of the "legal wrongs" recognized by American jurisdictions in the second-half of the twentieth century are similar to the torts just mentioned. They do not involve damage to bodily integrity or property, but are nevertheless interferences with valuable interests and are typically accompanied by emotional harm. These wrongs include eavesdropping, sexual harassment, disability discrimination, stalking, and certain instances of wrongful discharge. Legislatures have tended to create these causes of action within statutes that regulate by means other than private rights of action, so it is easy to overlook that these are, in some respects, contemporary torts.

E. Reconceiving Fear-of-Future-Disease Claims

The secondary legal quality apparatus affords a solution to the dilemma posed above. A defendant who negligently creates a heightened risk of a disease, under certain circumstances, may be negligently creating a substantial threat of serious illness. The *threateningness* of the circumstances that the defendant creates is a secondary legal quality of those circumstances. Its nature is constituted by the fact that a reasonable person would apprehend, from these circumstances, a substantial threat of contracting a serious disease in the future. This apprehension will typically go hand-in-

¹⁶⁶ Warren & Brandeis, *supra* note 8, at 196; see also Prosser, *supra* note 7, at 389.

¹⁶⁷ Prosser, *supra* note 7, at 389.

hand with fear and other emotional responses. But the threatening situation is itself an objective state of affairs whose qualities are not dependent upon a plaintiff's agency. Moreover, the interest protected, in the first instance, is an interest in being free from a certain kind of threat of disease rather than an interest in emotional tranquility per se. "The justification for permitting recovery for 'fear of' is that the plaintiff suffers, since '[l]ike the sword of Damocles, [plaintiff] knows not when it will fall.'"¹⁶⁸

A person who is wrongly subjected to a significant threat of a serious disease has been harmed in very important respects. A cloud has been placed over her life, and one can imagine that cloud of impending death intruding on her life significantly. Particularly if the risk is associated with an infectious and potentially fatal disease, the pall cast not only by the reduced prospects for normal life-expectancy, but also by the risk of infection, might undermine her self-confidence and cause her to alter significantly her plans, relationships, and activities. Moreover, just as false imprisonment causes a spatial confinement and thereby deprives one of an important form of liberty, so this cloud of impending disease constitutes a temporal confinement. The case law in this area illustrates the point. A Hawaii Supreme Court decision rendering distress based on exposure actionable reasoned that a plaintiff exposed to a pathogen or toxin may recover when "a reasonable person would foreseeably be unable to cope with the mental stress engendered by an actual, direct, imminent, and potentially life-endangering threat to his or her physical safety."¹⁶⁹ Accordingly, the court held that a cause of action exists if "the negligent behavior of a defendant subjects an individual to an actual, direct, imminent, and potentially life-endangering threat to his or her physical safety by virtue of exposure to HIV."¹⁷⁰

Working within standard analysis, the Hawaii court not surprisingly cast this cause of action as a claim for pure NIED. Yet it seems much more plausible to treat the claim like one for false imprisonment or defamation, with the plaintiff's emotional distress

¹⁶⁸ *Eagle-Picher Indus. v. Cox*, 481 So.2d 517, 527 (Fla. Dist. Ct. App. 1985) (quoting *Alley v. Charlotte Pipe & Foundry Co.*, 74 S.E. 885, 886 (N.C. 1912)).

¹⁶⁹ *John & Jane Roes, 1-100 v. FHP, Inc.*, 985 P.2d 661, 666 (Haw. 1999) (citing *Bramer v. Dotson*, 437 S.E.2d 773 (W. Va. 1993)).

¹⁷⁰ *Id.*

being compensated as a consequence of the invasion of an interest, rather than as constitutive of the interest itself. The plaintiff who is falsely imprisoned but remains unflapped still has a cause of action, for she has still been deprived of her important liberty interest. It seems plausible to say the same of a person who operates under the right sort of threat of disease, but remains unflapped.

It is a platitude in contemporary legal theory that what constitutes an injury is a normative matter, and our discussion on that point makes this platitude palpable. Undoubtedly it is tempting to treat what we have dubbed the “cloud” or “pall” associated with threat of illness as boiling down to probabilities: If the likelihood of the disease reaches a certain threshold, then it is actionable as a negligently created threat; otherwise it is not. But this is too facile. No doubt, percentage likelihood matters—the more remote the possibility, the less is it plausible to say that the defendant’s negligence is responsible (as opposed to the plaintiff) for the resulting pall. Indeed, we will suggest that the low percentage increase in risk in *Buckley* was important to the Supreme Court’s dismissal of the emotional distress claim brought there. But the question is not strictly factual. Whether the threat of disease and the interferences described constitute something significant enough to count as an “injury” seems to call for a more general account of what goods are needed for a life to go well or to be going well.¹⁷¹ In this respect, it resembles the question of what sorts of conditions constitute a “disability”: It is a normative question that hinges on an assessment of the capacities and activities that are important to a person’s life going well. We do not presume to address these issues. We merely note that the issue requires a normative judgment, made against the (usually unarticulated) conception of a life going reasonably well, and that it is not implausible that certain examples of “threat of a disease” could be viewed as such an interference as to support a cause of action for negligence.

¹⁷¹ Here courts might draw some guidance from modern disability law, and its conception of situations and conditions that substantially limit “major life activities.” See *Toyota Motor Mfg., Ky., Inc. v. Williams*, 534 U.S. 184 (2002) (discussing conditions that substantially limit major life activity); Samuel R. Bagenstos, Subordination, Stigma, and “Disability,” 86 Va. L. Rev. 397, 445-47 (2000) (arguing that the statutory definition of disability should incorporate a conception of what it is to lead a “normal” life).

The existence of a cognizable injury will also depend to a certain extent on how responsible the defendant's conduct was for creating the plaintiff's state of affairs. Without accepting as a given the Holmesian spectrum from intentionality to negligence to strict liability, that spectrum will be convenient for the present discussion. Indeed, as readers have probably noticed, many of the examples in our discussion, such as the idea of being under a cloud or threat, have involved actions that are usually brought in response to intentional wrongs: intentional assaults and false imprisonments, for example. Certainly, decisions in risk-of-disease cases such as *Potter* reflect a tendency among courts to be more willing to recognize a cause of action where the defendant's conduct arrives at, or even inches toward, intentional, knowing, or reckless maintenance of such a risk.¹⁷² Nor is this surprising. First, the monitoring costs required to avoid causing these injuries are much lower than they would be in negligence; the issue is not the sometimes delicate one of ensuring that one is acting with reasonable care, but rather requires only that the actor avoid purposefully or knowingly exposing another to a risk of injury. This means, in turn, that the diminution of the defendant's liberty is less substantial than it is under a regime of negligence-based liability. Second, the idea that the plaintiff's injury is something for which the defendant is responsible is much more plausible where the defendant has set out to cause this result, for the defendant's agency far overshadows that of the plaintiff in such scenarios. Third, although we generally shy away from any easy slippage from moral to legal, the intentional or reckless causing of such an interference is quite easily regarded as a moral wrong, and (as is indicated below) it is far from clear whether this is true in negligence.

As we have indicated in prior Sections and previous articles, the duty element in negligence law bears significant weight when plaintiffs seek to recover for injuries other than physical harm or property damage.¹⁷³ Pure economic harm and pure emotional harm are the two most notorious thickets within duty law, but the substantial threat of serious disease area should also be thorny. The threshold question in each of these areas is whether defendants are obligated

¹⁷² See *infra* note 222 and accompanying text.

¹⁷³ See Goldberg & Zipursky, *The Place of Duty*, *supra* note 6, at 668–72.

to be vigilant of the possibility of injuring plaintiff *in this respect*. The baseline or default is that, in the case of physical injury, there is an obligation to all others to take reasonable care not to cause such an injury through affirmative conduct. As for other kinds of harm, it is not obvious that there is such a general obligation, although these sorts of obligations attach in certain contexts and relationships. Moreover, as we have argued, and as Mark Geistfeld has eloquently elaborated from a different point of view in this Volume, part of the role played by the notion of obligation within our moral psychology is that it is important if duty is to function at all that it prioritize certain interests above others.¹⁷⁴ This prioritization is not consistent with finding a general obligation to be vigilant of all kinds of harm.

Of course, negligence is a broader tort than false imprisonment, assault, defamation, and invasion of privacy. An interesting question is whether the more carefully drawn contours of those other torts permit them to protect special sorts of interests without unduly cutting into the prioritization of physical well-being that negligence law requires. To put the point the other way around (and probably more accurately, historically), it is plausible that negligence has been able to develop as a broad supplement to specific torts only because the duty of due care as a general obligation has been limited to physical injuries caused by misfeasance. Whichever way the point is put, however, negligence law has progressed by pushing duty beyond misfeasance and physical harm. Courts have generally tried to do this incrementally, in a manner that preserves the notion of obligation, meshes plausibly with social and professional norms, does not interfere greatly with other institutions or branches of law, and remains manageable within the judicial system.¹⁷⁵

The questions at hand are therefore whether and when defendants should be viewed by courts as having a duty to others to be vigilant not to cause them to be under a cloud of serious disease. As argued in prior Sections, it is not enough to argue that there is a

¹⁷⁴ See Geistfeld, *supra* note 69; Goldberg & Zipursky, *supra* note 4, at 1830–39.

¹⁷⁵ See Goldberg & Zipursky, *The Place of Duty*, *supra* note 6, at 740 (describing considerations relevant to duty analysis).

duty not to cause the target disease.¹⁷⁶ It is, to speak metaphorically, the injury of the "cloud" or "threat" for which we envision plaintiff recovering. Therefore, the duty established must be a duty to be vigilant not to cause this cloud or threat.

Who has such a duty, and to whom? To some extent, this is an issue that jurisdictions must answer on a case by case basis, but this—rather than questions of floodgates and fraud—is the question courts should be asking. Some cases also present themselves as much stronger than others. A hospital—particularly one that treats a significant number of HIV patients—seems to have a duty to its patients and to its sanitary workers to be vigilant not to expose them to the risk of a needle stick that may expose one of them to HIV. On the other hand, it is less obvious that a used car dealer owes a duty to future users to make sure they are not exposed to HIV from debris hidden inside the car.¹⁷⁷ An intermediate, and very difficult question, is whether the duty to be vigilant of causing a threat of HIV infection involves a duty not to cause *actual* exposure to HIV or a duty not to cause *possible* exposure through a medically possible means of transmission. Perhaps the best course is to follow courts that have availed themselves of burden-shifting to ease this difficult dilemma.¹⁷⁸

Although the toxic exposure cases present equally challenging questions, our aim here is merely to find a sufficiently rich and principled framework for asking the questions, not to prove that one particular answer ought to prevail in every jurisdiction. For example, a powerful argument can be made that an industry producing hazardous waste or radiation has a duty to the inhabitants of

¹⁷⁶ See *Boyles v. Kerr*, 855 S.W.2d 593 (Tex. 1993) (holding that an independent duty is required).

¹⁷⁷ See, e.g., *Cole v. D.J. Quirk, Inc.*, 2001 Mass. App. Div. 139, 142 (2001) (adopting actual exposure rule to reject claim of persons who were stuck by surgical tweezers in used car).

¹⁷⁸ See, e.g., *S. Cent. Reg'l Med. Ctr. v. Pickering*, 749 So. 2d 95, 102 (Miss. 1999) (shifting the burden of proof against actual exposure to the defendant after the plaintiff had shown possible exposure in a diabetic lancet case). Of course, fear-of-AIDS claims must also be limited by the so-called "window of anxiety"—the period between exposure and the time at which medical tests can rule out infection. *Id.* at 103. As faster, reliable tests become available, the scope of liability in this area will shrink accordingly.

nearby towns not to create a substantial threat of cancer.¹⁷⁹ It will be a difficult question, however, what constitutes such a threat; similarly, it will be a difficult question whether defense contractors have a parallel duty not to cause the threat of serious disease to all those who might be affected by their weaponry, and to their descendants.

Perhaps the best way to bring our analysis of emotional distress claims to closure is to return to the decision with which this Article started: *Buckley*. The plaintiff's case seems initially compelling because Metro-North's conduct was negligent, and perhaps worse. And yet FELA provides private rights of action for injured plaintiffs; it is a tort statute, not a regulatory law. The Court rightly looked beyond Metro-North's wrongful conduct to see whether Buckley was genuinely seeking redress for an injury, that is, whether Metro-North's wrong was realized in an injury.

Buckley argued that his exposure to asbestos was itself an impact, which supported an action for personal injury.¹⁸⁰ On reflection, it should be clear that this "impact" constituted what we have called an "intermediate harm." To be sure, it stopped Metro-North from asserting that its negligence was merely "in the air"; it was literally *in* Buckley himself. But it was not cancer. Even if it had the potential to become that disease, it had not yet ripened. For reasons both pragmatic and principled, Justice Breyer rightly noted that it was inadvisable and alien to negligence law to treat such unripened injuries as a basis for a right of action.¹⁸¹ Moreover, he rightly resisted the temptation to adopt an even more minimal, and arbitrary, rendition of the impact rule.¹⁸²

¹⁷⁹ See *Day v. NLO*, 851 F. Supp. 869 (S.D. Ohio 1994) (permitting recovery for fear of cancer).

¹⁸⁰ *Buckley*, 521 U.S. at 428.

¹⁸¹ *Id.* at 430.

¹⁸² The petitioner's argument in *Norfolk & Western Railway Co. v. Ayers*, 122 S.Ct. 1434 (2002) (mem.), can be characterized in at least two ways. On the one hand, it might be construed as an invitation to the Court to abandon the predicate/parasitic distinction. In our view, that invitation should be declined. The Court has repeatedly maintained that FELA must be interpreted in light of common law, and the distinction is a sound and well-settled feature of that law.

On the other hand, it may pose the question of whether the form of asbestosis alleged is in fact so minimal as to be nothing more than an "impact" of the sort *Buckley* found insufficient to support a claim for fear of future disease. We take no position on whether this latter characterization is correct.

The question therefore becomes whether there is a sense in which Metro-North, in negligently exposing Buckley to asbestos and thereby enhancing the likelihood that he would get cancer, was causing not only an intermediate harm, but an ultimate harm; not simply a realizable tort but a realized one. Buckley's claim that he was emotionally injured was an effort to make exactly such a claim. Citing reasons of credibility and floodgates, the Court rejected Buckley's claim, and with it, the idea of liability for exposure alone.¹⁸³

It is easy to see that Buckley's own claim of emotional harm was weak. A heavy smoker who never sought psychiatric or psychological treatment, Buckley's fear of cancer claim was tenuous on the facts. Even Justice Ginsburg recognized this, and (on that point) concurred in the denial of relief.¹⁸⁴ Yet, as the dissent pointed out, it is far from clear why there ought to be a denial of relief across the board for such claims simply because Buckley's own claim was weak.¹⁸⁵ And in the face of a wealthy and negligent defendant, and a recent decision in *Gotshall* expressing openness in principle to damages for emotional harm under FELA,¹⁸⁶ the Court's wholesale refusal to recognize a cause of action rang hollow.

We have argued above that unrealized toxic-exposure cases should be viewed in a different light—not as *fear-of-disease* cases but as objectively based *threat-of-disease* cases. This account explains both why *Buckley* was in some ways a close case and also why it was ultimately correctly decided on this issue. *Buckley* was close because on a key issue—duty—the plaintiff's position was strong. An employer like Metro-North, given the state of medical knowledge in the 1980s, very plausibly owed a duty not only to be vigilant against causing its employees certain serious diseases, but also a duty to be careful not to cause a substantial threat of those diseases. In other words, the objective circumstances of being in working conditions that cause one to live under the cloud of impending disease are circumstances an employer plausibly has an obligation to take care not to cause to its employees. Metro-North

¹⁸³ *Buckley*, 521 U.S. at 433–34.

¹⁸⁴ *Id.* at 444.

¹⁸⁵ *Id.* at 445, 454 (Ginsburg, J., dissenting).

¹⁸⁶ *Id.*

took no such care. To that extent, Buckley's claim—though novel—appears strong.

On the other hand, as Justice Breyer seems to have recognized, Buckley's particular claim was quite weak.¹⁸⁷ This is not simply because he did not seem very anxious about his risk of cancer,¹⁸⁸ but because the "threat of cancer" to which Buckley was exposed simply was not very significant. Although Metro-North's conduct may have been wrongful, and Buckley should not have been placed in the circumstances in which he was placed, Buckley was not placed under a serious threat by this conduct. Indeed, his own experts testified that the increased risk of cancer was from 1% to 5% (Justice Breyer supposed, hypothetically, that Buckley's risk of cancer moved from 23% to 28%).¹⁸⁹ While it is not always easy to find the right numbers in a case such as this, there is little plausibility to the assertion that a marginal increase of a few percentage points in an already substantial risk of cancer—while surely of concern to the person subjected to the increase—converts a situation that was psychologically tolerable into one in which the person finds herself under "Damocles' sword." At the end of the day, there was no injury in *Buckley*—no tort at all. Thus, even if Buckley had presented evidence that he was significantly distressed, he should not have been permitted recovery because that distress could not fairly be imputed to the circumstances created by the defendant's negligent conduct. To impose liability for such a small increment of increased risk would truly be to regulate under the guise of tort, rather than to provide recourse for a person who has suffered a realized wrong.

IV. MEDICAL MONITORING CLAIMS

Suits seeking reimbursement for medical monitoring expenses provide yet another way to frame claims in tort for heightened risk of future physical injury or illness. Here, again, we have to be careful about what sort of claims we have in mind. Medical monitoring suits refer to actions in which the cost of medical monitoring is the only cognizable injury alleged by the plaintiff. Thus, we do not ad-

¹⁸⁷ Id. at 427.

¹⁸⁸ Id. at 428 (noting the trial court's finding that Buckley presented little evidence of real emotional injury).

¹⁸⁹ Id. at 427, 435.

dress suits seeking reimbursement for medical monitoring expenses as damages parasitic on some other injury, such as a physical injury. The latter suits raise issues in the law of remedies, but typically there is no obstacle (beyond issues of proof) to recovery of medical monitoring expenses. For example, a plaintiff whose heart valve is damaged as a result of ingesting a defectively or negligently designed product would ordinarily be entitled to recover the expected cost of annual tests to monitor the deterioration of the valve.¹⁹⁰ The same analysis would presumably apply to cases falling within the "penumbral" category described in Section III.B. If a negligence plaintiff can establish that she is actually under a threat of serious illness as a proximate result of defendant's carelessness, then, in principle, she ought—like the plaintiff with the damaged heart valve—to be able to recover for the expected medical monitoring expenses necessitated by the threat.

In short, medical monitoring claims are akin to claims for enhanced risks of future injury, except that they seek to recover the expected cost of preventive medical treatment necessitated by another's wrongful conduct rather than a percentage of the value of the ultimate harm expected to flow from that conduct. We have said earlier that the common law governing negligence and products liability does not ordinarily treat increased risk as an injury in and of itself. Yet clearly many medical monitoring plaintiffs are prevailing in the courts. Thus, we need to consider whether they are entitled to prevail and, if so, on what grounds.

A. Uncertainty over the Nature of Medical Monitoring Claims

Genuine medical monitoring claims have become an important part of the torts landscape. Presumably, this is a function of a heightened sensitivity to both the presence of toxins in everyday

¹⁹⁰ See Restatement (Second) of Torts § 924(c) (1979) ("One whose interests of personality have been tortiously invaded is entitled to recover damages for . . . prospective . . . (c) reasonable medical and other expenses . . ."). To draw a sharp contrast, consider a case in which C Corp. carelessly operates a factory so as to introduce vast amounts of a known carcinogen into a neighboring town's water supply. If we suppose that the residents' exposure to the carcinogen increased their risk of contracting a form of cancer to 1 in 10 from a background risk of 1 in 1,000, the residents might fairly complain of being under the sort of threat that would support recovery of damages for fear of future illness.

life and to the perceived link between those toxins and certain injuries. Equally, it may reflect the problems of proof that have given rise to loss-of-chance and enhanced risk claims. Many toxic tort plaintiffs cannot establish as more probable than not a general causal link between a given product or course of conduct and a given injury or illness. This information gap usually defeats claims for physical injury. It may even defeat claims for emotional distress, since the absence of evidence of a causal link might render the emotional distress unreasonable as a matter of law. Medical monitoring claims do not, in principle, require plaintiffs to establish that they are under a significant threat of injury or illness; rather, they only require plaintiffs to establish that a reasonable regime of preventive medicine would include periodic tests to detect the onset of the illness.¹⁹¹

Pure medical monitoring claims have also arisen in part because of particular aspects of the law of civil procedure. Medical monitoring claims are often brought as class actions on behalf of all persons who ingested a particular medication or were exposed to a particular toxin. The traditional mass tort class action attempts to certify a class under Federal Rule of Civil Procedure 23(b)(3), which requires that common questions of law and fact "predominate" over individualized issues.¹⁹² Courts of Appeals and the Supreme Court have, however, cast suspicion on the invocation of Rule 23(b)(3) for mass tort class actions, given that common questions (for example, whether a given product was defectively designed) are often overwhelmed by individual issues such as compliance with statutes of limitations, causation of particular plaintiffs' injuries, damages, comparative fault, and assumption of risk.¹⁹³ In response, some lawyers have sought class certification for medical monitoring claims under Rule 23(b)(2), which authorizes class actions on the ground that the defendant has "acted or refused to act on grounds generally applicable to the class, thereby making appropriate final injunctive relief or corresponding declaratory relief

¹⁹¹ See, e.g., *Ayers v. Jackson Township*, 525 A.2d 287, 304 (N.J. 1987) (distinguishing claims for enhanced risk from claims for medical monitoring).

¹⁹² Fed. R. Civ. P. 23(b)(3).

¹⁹³ See *Amchem Prods. v. Windsor*, 521 U.S. 591, 624 (1997) (quoting *Georgine v. Amchem Products*, 83 F.3d 610, 626 (3d Cir. 1996)).

with respect to the class as a whole.”¹⁹⁴ The theory behind this sort of claim is that such suits are not seeking damage payments per se—which would bar certification under 23(b)(2)—but rather are asking the court to enjoin a defendant who has put a class of plaintiffs at heightened risk of developing a disease to establish a fund or program by which class members can pay for medical monitoring of their risk.

Friends for All Children v. Lockheed Aircraft Corp. is sometimes cited as announcing the arrival of medical monitoring as a cognizable harm.¹⁹⁵ There, the court upheld a district court decision ordering the defendant to create a fund to pay for neurologic testing of children placed at risk of injury by the decompression and crash of the defendant’s plane.¹⁹⁶ Subsequent to *Lockheed*, the New Jersey and California Supreme Courts endorsed medical monitoring claims in cases alleging heightened risk of cancer due to localized environmental contamination. These decisions, *Ayers v. Jackson Township*¹⁹⁷ and *Potter v. Firestone Tire & Rubber Co.*,¹⁹⁸ have emerged as the leading authorities in the area.

To say that medical monitoring claims are important and have been recognized by influential state supreme courts is not to say that courts have settled on an adequate understanding of these claims. In fact, these claims have raised many complex issues of both substantive and procedural law. What degree of heightened risk is required before a claim will be recognized? Is this determination medical or legal? Does class certification under Rule 23(b)(2) permit an end-run around the strictures of Rule 23(b)(3) by allowing plaintiffs first to obtain certification for purposes of equitable relief, then to secure a cash settlement with a value that is affected by the certification decision? Under the rules against

¹⁹⁴ Fed. R. Civ. P. 23(b)(2).

¹⁹⁵ 746 F.2d 816 (D.C. Cir. 1984).

¹⁹⁶ *Id.*; see generally Arvin Maskin et al., Medical Monitoring: A Viable Remedy for Deserving Plaintiffs or Tort Law’s Most Expensive Consolation Prize?, 27 Wm. Mitchell L. Rev. 521 (2000) (describing the relevant case law).

¹⁹⁷ 525 A.2d 287 (N.J. 1987).

¹⁹⁸ 863 P.2d 795 (Cal. 1993).

claim-splitting, does recovery of medical monitoring costs preclude a later suit for realized injuries such as cancer?¹⁹⁹

We are not in a position to take on these issues in the present discussion. Rather, as indicated at the outset, we aspire to isolate the substance of typical medical monitoring causes of actions in the hope that doing so will help resolve these and other issues swirling around medical monitoring suits. Courts, understandably, have not yet settled on a satisfactory account of the nature of medical monitoring claims. Some have simply denied the viability of medical monitoring suits as presenting cognizable claims in tort.²⁰⁰ The deep ambivalence expressed by the Supreme Court in *Buckley*, however, is more illustrative of the problems of characterization that medical monitoring suits pose.

As noted at the outset, the *Buckley* majority—concerned about the sheer amount of potential liability, as well as the problem of giving de facto priority to claims by the as-yet-uninjured over claims by the injured—showed a strong inclination to reject such claims under FEHA.²⁰¹ Yet, “enigmatically,” it refused to slam the door on them. Instead the Court left open the possibility that such claims would be actionable if a court entertaining the claims ordered the right manner of remedy, namely the establishment of a fund from which patients could obtain ongoing reimbursement for medical monitoring expenses rather than a lump-sum payment.²⁰²

In its ambivalence, the *Buckley* majority never really came to grips with the nature of the plaintiff’s medical monitoring claim. Repeatedly, it characterized the claim as one for negligence caus-

¹⁹⁹ See generally McCall, *supra* note 66 (noting conflicting decisions on this question and arguing against the assignment of preclusive effect to successful claims for medical monitoring).

²⁰⁰ These decisions seem to have been rendered predominantly by federal courts that, in exercising diversity jurisdiction, are reluctant to develop state tort law. See, e.g., *Ball v. Joy Techs.*, 958 F.2d 36, 39 (4th Cir. 1991) (requiring proof of a predicate physical injury); *Carroll v. Litton Sys.*, No. B-C-88-253, 1990 U.S. Dist. LEXIS 16833, at *148-49 (W.D.N.C. Oct. 29, 1990) (denying claim based on increased risk of disease). For a collection of decisions recognizing and rejecting medical monitoring claims, see Allan L. Schwartz, Annotation, Recovery of Damages for Expense of Medical Monitoring to Detect or Prevent Future Disease or Condition, 17 A.L.R.5th 327 (1994).

²⁰¹ See *supra* text accompanying notes 17-25.

²⁰² See *supra* text accompanying notes 21-25.

ing economic loss.²⁰³ As mentioned at the outset, to frame the claim this way is to link medical monitoring claims such as Buckley's to actions such as those against accountants for negligent misrepresentation. Yet this linkage is not very satisfactory. In the first place, Buckley's claim was not for economic losses already experienced, but for future economic losses he expected to incur.²⁰⁴ At a minimum, the suit was premature, at least until Buckley had to pay for his first set of medical tests. Even then, it is unclear whether an economic injury of this nature would support recovery of expected future expenses as parasitic damages.²⁰⁵

More fundamentally, we doubt that claims like Buckley's are viable as claims for pure economic loss. No jurisdiction recognizes a general duty to take care to avoid causing economic loss to others.²⁰⁶ Indeed, the dominant rule remains the rule of *Robins Dry Dock and Repair Co. v. Flint*: In the absence of damage to physical property, economic loss is not recoverable absent privity between the defendant and the plaintiff or some other special relationship that attests to the defendant have undertaken to be vigilant of the plaintiff's intangible economic interests.²⁰⁷ While it is true that Metro-North and Buckley maintained an employer-employee relationship, we doubt that this is the sort of relationship to which a duty to be vigilant of economic well-being attaches. First, management is often under obligations to owners that preclude obligations to be vigilant of economic losses to employees. (Consider, for example, a decision to approve a merger or close a plant, which will foreseeably cause the loss of jobs.) More to the point, even when not responding to such obligations, managers retain discretion to make decisions without regard to their economic impact on employees. Absent enforceable promises to the contrary, and assum-

²⁰³ *Buckley*, 521 U.S. at 438-40.

²⁰⁴ *Id.* at 438.

²⁰⁵ See Restatement (Second) of Torts § 924(c) (1979) (limiting this sort of recovery to "one whose interests of personality have been . . . invaded").

²⁰⁶ See Dobbs, *supra* note 19, § 452, at 1282 ("When commercial or economic harm stands alone, divorced from injury to person or property, courts have not imposed a general duty of reasonable care. They have instead imposed a duty . . . mainly when the parties are in a special relationship and the defendant has implicitly undertaken such a duty."). Ironically, then-Judge Breyer is the author of a leading opinion on this issue. See *Barber Lines A/S v. M/V Donau Marie*, 764 F.2d 50, 54-55 (1st Cir. 1985).

²⁰⁷ See *Robins Dry Dock & Repair Co. v. Flint*, 275 U.S. 303, 308-09 (1927).

ing they have not made decisions in a discriminatory or otherwise unlawful manner, if the managers of a firm decide to move their operations across town so as to force their employees to endure longer and more expensive commutes or to pay higher taxes, there would be no cause of action even if the decision was careless as to the employees' economic interests. The same would hold true if they were to shut down the office child care center, thus forcing certain members of the workforce to quit or work reduced hours. Likewise, if management decided, for business reasons, to switch the company's health plan from one insurer to another, thereby increasing employees' deductibles and reducing their coverage, no cause of action in negligence would accrue.

To point to these examples is not to say that employers owe no obligations to their employees. Employers owe employees what they have promised them. They also owe duties set by federal and state statutes and regulations with respect to nondiscrimination, workplace safety, etc. In addition (discussed below), employers owe employees various duties in tort. For example, as *Buckley* acknowledges, employers sometimes owe employees a duty to avoid causing situations that generate emotional distress.²⁰⁸ Rather, all we aim to have established is the implausibility of the proposition that employers such as Metro-North have an obligation to be vigilant not to cause intangible economic harm to their employees. If we are correct about this, it tells us that medical monitoring claims are not on sound footing when cast as claims for economic loss, as they were in *Buckley*.²⁰⁹ If they are to be recognized, they must be recognized on some other doctrinal ground. Otherwise, they revert to the status of claims for pure enhanced risk of physical injury, and invoke the general rule in tort law that heightened risk, in and of itself, is not a cognizable harm.²¹⁰

Apart from the primary duty question, there is a second anomalous feature of *Buckley*, hammered home by Justice Ginsburg in her dissent, concerning the majority's apparent conflation of right and remedy. Previous decisions on which *Buckley* relied, including *Ayers* and *Potter*, had suggested that a lump-sum payment was a

²⁰⁸ See supra text accompanying notes 16–17.

²⁰⁹ See supra text accompanying notes 65–71 (discussing the general absence of a duty to avoid causing intermediate harms).

²¹⁰ See supra text accompanying note 10.

less desirable form of relief than some sort of structured payment scheme. Neither, however, insisted on this form of remedy. In *Ayers*, the New Jersey Supreme Court declined to impose that remedy in the case at hand, suggesting instead that it should be the "general rule" in future medical monitoring suits against public entities.²¹¹ *Potter* merely noted its advantages.²¹² In this respect, both are consistent with the conventional notion, expressed by Justice Ginsburg, that the issue as to whether medical monitoring is a viable cause of action is wholly distinct from the question of how to structure the remedy.²¹³ The former is determined by the legal rules setting out the elements of the cause of action; the latter is typically left to the discretion of the trial court. *Buckley*, by contrast, seems to build the form of relief into the substance of the cause of action: A medical monitoring claim is viable only if the remedy consists of the establishment of a fund rather than the payment of lump-sum damages.

In sum, the oddity of *Buckley* is not its decision to restrict claims for medical monitoring. As noted earlier, the majority was motivated by important considerations of policy. Rather, the oddity resides in the way in which the Court's policy concerns were translated into tort law via FELA. By focusing on the economic injury and the issue of remedy, the Court seems to have rendered a decision that does not "carve at the joint." Instead, it appears to achieve an ad hoc compromise between *principle* (which tells us that *Buckley* has suffered a cognizable wrong in the form of negligent infliction of economic loss) and *policy* (which tells us that the *Buckleys* of the world should not recover because of the threat of too much liability and the prioritization problem). Like the old privacy decisions Dean Prosser studied, *Buckley* indicates that the cause of action for medical monitoring is as yet inadequately realized.²¹⁴

²¹¹ *Ayers*, 525 A.2d at 314-15.

²¹² *Potter*, 863 P.2d at 825 n.28.

²¹³ *Buckley*, 521 U.S. at 455-56 (Ginsburg, J., dissenting).

²¹⁴ See Prosser, *supra* note 7, at 422-23.

B. Medical Monitoring and the Duty to Aid the Imperiled

We have expressed doubts concerning the *Buckley* majority's characterization of a medical monitoring cause of action as a duty to avoid causing economic loss. We have also noted that it seems to conflate right and remedy, but that this conflation was motivated by legitimate concerns about the scope of liability and the priority to be accorded claims of actual injury versus claims for monitoring expenses. Thus, to this point, we might fairly be taken to be arguing that the right realization to have about medical monitoring claims is that courts ought never to recognize them. That, however, is not our position. In fact, we are inclined to think that the *Buckley* majority was right to link the form of the remedy to the underlying right of action, although it did not adequately explain why. Once this link is better understood, the true doctrinal basis for a medical monitoring claim emerges.

The key here is to realize that defendants who are ordered to underwrite medical tests for those whom they have put at risk of physical injury should not be understood as providing a compensatory payment for injuries already realized, economic or otherwise. This is because, at least on facts such as *Buckley's*, they have not yet committed a tort. Of course, when one sees money changing hands in the context of a tort suit, one naturally thinks in terms of compensation. Ordinarily, payment is made in satisfaction of what is sometimes described as a *secondary* duty of repair. It is described as secondary, because it attaches as the result of a violation of a *primary* duty not to cause an injury through wrongful conduct.²¹⁵ In this instance, however, appearances are misleading. Payment in a tort context need not always be in satisfaction of a secondary duty to compensate for a completed wrong. Rather, payment might also be a way of performing a certain kind of primary duty owed by a defendant to a plaintiff. And this is more likely to be the case when payment comes in the form of a fund issuing periodic payments rather than a lump-sum transfer: The mode of payment bespeaks an ongoing obligation, rather than a one time satisfaction of a debt owed in light of a wrong already done.

²¹⁵ Jules L. Coleman, *Tort Law and the Demands of Corrective Justice*, 67 *Ind. L.J.* 349, 367 (1992).

Recasting payment of medical monitoring expenses as the performance of a primary duty of conduct rather than the fulfillment of a secondary duty to compensate might seem to get us out of the frying pan and into the fire. If such payments constitute the performance of a primary duty, then the duty in question must be one of affirmative aid, as opposed to a duty to take care to avoid causing injury. Notoriously, the common law of tort is, and has long been, quite stingy about recognizing such affirmative duties. There is no general duty to protect or rescue through affirmative conduct, as opposed to a duty to refrain from acts that injure a person.²¹⁶ Why, given the paucity of affirmative duties recognized by the common law, should we suppose that we have found an adequate ground for medical monitoring claims by reconceiving them in terms of an affirmative primary duty?

The answer lies in the fact that one of the well-recognized exceptions to the general rule against affirmative duties is the duty owed by one who has created a dangerous condition that renders another in peril and hence in need of affirmative aid.²¹⁷ A storeowner who would otherwise not owe an affirmative duty of reasonable care to a person in the store may incur a duty to take steps to aid the person if she is injured by an instrumentality or employee of the store.²¹⁸ Furthermore, if a person buys narcotics for another, she has a duty to assist the other if he requires medical attention as a result of using the drugs.²¹⁹ Or suppose *D*, driving his car at night, runs another driver, *P*, off the road without causing injury to *P* or his car. *D* then drives off without attempting to help *P* by, for example, shining his headlights in *P*'s direction so that *P* can assess the situation. *P* falls down a steep embankment and is injured, and can show that he would not have fallen had *D* shone his headlights on the accident scene. *D* will be liable to *P* because, once having placed him in peril, he was obligated to take steps to protect *P* from the peril.

Metro-North was in an analogous position to these defendants vis-à-vis its employee Buckley. By causing Buckley to be exposed

²¹⁶ Dobbs, *supra* note 19, § 314, at 853.

²¹⁷ *Id.* § 316, at 856–57.

²¹⁸ *L.S. Ayres & Co. v. Hicks*, 40 N.E.2d 334, 337–38 (Ind. 1942).

²¹⁹ *Heffern v. Perry*, No. 420713, 2000 Conn. Super. LEXIS 2708, at *1 (Conn. Super. Ct. Oct. 11, 2000).

to substantial amounts of airborne asbestos in the workplace, it subjected him to physical peril. At that point, the company incurred a primary duty to Buckley to take reasonable measures to help him avoid realizing the peril. By paying for the medical monitoring necessitated by the exposure, Metro-North fulfills that duty. Providing such monitoring, or the money to fund such monitoring, is not a symbolic, retrospective form of redress for an (economic) injury done. It is a forward-looking act that Metro-North is obligated to perform in light of having created the situation in which such an act is needed to minimize the risk of injury. Thus, Buckley's suit, properly conceived, was not seeking redress for a completed wrong, but an injunction ordering Metro-North to attend to a duty that, if left unattended and if later linked to a genuinely cognizable harm, would then (and only then) ripen into a tort.

Our reconstruction of medical monitoring claims in terms of seeking to enjoin performance of a primary duty rather than seeking compensation for a completed wrong better captures three aspects that have thus far remained unilluminated by judicial analysis. First, it explains, at the level of doctrine, the plausibility of these claims. Cast as compensatory payments for injuries already suffered, medical monitoring awards are styled as compensation for either economic losses or heightened risk of future physical injury. Yet, the latter is ordinarily non-cognizable in negligence, and the former is forced and implausible in attributing to defendants such as employers a duty to be vigilant not to cause economic losses to their employees. By contrast, the description of medical monitoring payments as fulfilling a primary duty provides an accurate description of the nature of the wrong and the nature of the duty in these cases. It is plausible for a court to say to Metro-North: "You have done wrong by carelessly imperiling your employee, now it is your responsibility to take steps to reduce the likelihood that this peril is realized, or to mitigate the severity of the harm experienced if it is realized."

Second, our reconstruction renders intelligible the otherwise mysterious linkage drawn in *Buckley* between right and remedy. *Buckley* itself treated structured payments as an ad hoc limitation designed to serve the policy goals of reducing overall liability for medical monitoring claims. By contrast, on the present account, it is precisely because the duty owed by the defendant is the duty to

provide for monitoring—rather than the secondary duty to pay compensation—that lump-sum payment is inappropriate as compared to periodic payments as monitoring costs are incurred. Indeed, scheduled payments appear, in this light, to be much more sensible than a one-time, lump-sum payment.

Third, our account makes sense of the efforts to plead medical monitoring claims as class actions under Rule 23(b)(2). Class members in these cases are asking for an equitable remedy, rather than a damage award. By way of duty determination, they are asking the court to determine that the defendant has put them in peril and therefore has incurred a duty to take care to minimize their peril. By way of relief, they ask the court to order the defendant to heed that duty by paying for medical monitoring on an ongoing basis.

Our “solution” to the puzzle of medical monitoring claims may strike some readers as contrived. In fact, we think it integrates these claims into the existing fabric of tort law quite well. We acknowledge, however, that the fit may seem awkward in three respects. First, in terms of substantive doctrine, the duty to aid those whom one has imperiled attaches even when the imperilment is innocent, rather than wrongful. Thus, in the foregoing example of the driver, *D*’s duty to help *P* would attach even if *D* was driving with all reasonable care when he ran *P* off the road.²²⁰ By contrast, suits like those in *Ayers*, *Buckley*, and *Potter* sometimes allege grave wrongdoing, and, more importantly, the courts seem to take it for granted that there must be a showing of wrongdoing (at least negligence) before medical monitoring costs will be awarded.

Second, in terms of procedure, affirmative duties such as the duty to aid those whom one has imperiled are almost always enforced after the fact, by means of compensatory payments for injuries flowing from the imperilment. One cannot typically sue for a positive injunction ordering a rescue effort to be undertaken. (Of course in many, if not most cases, there is only a momentary lapse in time between imperilment and injury, so the question does not usually arise.) Yet, medical monitoring claims seek and obtain equitable relief prior to injury. Taken together, these differences in

²²⁰ See Dobbs, *supra* note 19, § 316, at 856–57 (noting that the duty to aid applies even in cases where peril is innocently created).

substance and procedure pose a third problem, namely the potentially broad reach of our solution to the doctrinal conundrum of monitoring claims. If the innocent creation of peril to another is sufficient to obtain a court order requiring payments necessary to perform the duty to minimize the peril, are we not in danger of inviting the aggregate levels of liability, as well as the undesirable prioritization of claims, that Justice Breyer rightly worried about in *Buckley*?

The short answer to the last question is "yes." The duty to aid those whom one has imperiled would impose substantial burdens and might well, in limited-fund cases, steer resources to the "merely" imperiled at the expense of those in whom the peril has been realized. One might imagine, for example, a class action on behalf of all non-smokers exposed to substantial levels of tobacco smoke (such as bartenders or waitstaff) against manufacturers of cigarettes and cigars for the cost of monitoring for emphysema or lung cancer. Not only would the expense be enormous, given the sheer number of plaintiffs, it would likely divert money from those who have developed those diseases.

Still, recognition of the potential overbreadth problem actually helps generate answers to the other two puzzles we have just noted: the problem of the innocent imperiler and the availability of pre-injury relief. The relief sought by medical monitoring plaintiffs is equitable in nature. As a matter almost of definition, courts retain broad discretion within their equitable jurisdiction to order what is just. Recall the earlier *Vaughn v. Menlove* hypothetical. A judge faced with a request from Menlove to enjoin Vaughn to eliminate the dangerously constructed haystack ordinarily will not issue the injunction unless the haystack is found to constitute a nuisance (which, in the facts of the hypothetical, it is not). The law, a court likely would reason, contains an adequate remedy in the form of the action for negligence and the payment of after-the-fact compensatory damages.²²¹ Likewise, in most other cases of imper-

²²¹ Many scholars, most prominently Professor Laycock, have suggested that, as employed by the courts, the availability of an adequate remedy at law does not constitute a ground in and of itself to deny equitable relief. See Douglas Laycock, *The Death of the Irreparable Injury Rule 4* (1991). Insofar as this is meant to suggest only that judges do not enforce a blanket prohibition against pre-injury equitable relief in tort cases, we agree. There seems little doubt, however, that courts do not routinely grant

ilment, the judicial default should be the denial of injunctive relief in favor of legal remedies after the peril is realized. The potentially substantial burdens that attach to affirmative duties to minimize perils, coupled with the fact that the peril has not yet been realized, and that, if realized, will support a cause of action for the injury, work together to disfavor, in the ordinary case, the granting of injunctive relief.

The balance of equities may change, however, in particular cases, so as to favor the grant of injunctive relief. If the imperiled plaintiff can establish that the defendant has not merely innocently imperiled him, but instead has imperiled him by virtue of wrongful conduct toward the plaintiff, the defendant's argument against being enjoined is weakened while the plaintiff's argument for injunctive relief is bolstered. Indeed, in this regard, we note that in *Ayers*, *Buckley*, and *Potter*, there were allegations not simply of negligence, but of gross negligence bordering on recklessness, which no doubt aided the case for equitable relief.²²² Likewise, if the defendant and the imperiled plaintiff are in a pre-existing relationship within which the defendant already owes the plaintiff certain obligations to protect her well-being, the equities again shift toward the plaintiff. Finally, if the relationship is one that is already subject to significant legal regulation designed to protect the plaintiff from exposure to the risks in question—as workplace hazards are regulated by the Occupational Safety & Health Administration ("OSHA") and the discharge of contaminants is regulated by the Environmental Protection Agency ("EPA") and its state equivalents—there is yet another reason to favor an injunctive remedy. In

such relief in tort cases, and that one of their reasons for doing so is because a post-injury action may be brought. See 2 Dan B. Dobbs, *Law of Remedies: Damages-Equity-Restitution* § 8.10, at 535 (2d ed. 1993) (noting the traditional reluctance of courts to grant injunctive relief in connection with personal injury claims).

²²² See *Buckley*, 521 U.S. at 443 (conceding the accuracy of Justice Ginsburg's dissent in describing plaintiff being sympathetic for having suffered a wrong at the hands of his employer); see *id.* at 446-47 (Ginsburg, J., dissenting) (describing the facts and noting the state agency's conclusion that the employer had "seriously disregarded the health and safety of its workers," and that the employer's failings were "particularly egregious"); *Potter*, 863 P.2d at 818 (relieving the plaintiffs of an obligation to prove that they were more likely than not to develop illness due to an exposure caused by the defendant because of the defendant's "conscious disregard of the rights and safety of others"); *Ayers*, 525 A.2d at 292 (noting ample evidence that the town operated its landfill in a "palpably unreasonable" manner).

this regard, we note that both Justice Breyer for the majority and Justice Ginsburg in dissent pointed out that OSHA already required that many employers provide safe workplaces, and even that they pay for medical monitoring of employees exposed to toxins.²²³

In sum, our characterization of claims for medical monitoring expenses—claims of the sort upheld in *Ayers* and perhaps tolerated in *Buckley*—as actions by imperiled plaintiffs seeking equitable relief to enforce the defendants' primary duty to minimize the peril does not "prove too much." Instead, it can explain why relief is sometimes granted, while still being entirely consistent with a high degree of judicial caution in the expansion of this particular remedy. Our view, in other words, is completely consistent with a default rule against the recognition of such claims. In situations where an actor is already subject to regulations to protect certain parties from exposure to toxins, and where the actor engages in wrongful conduct that rises to the level of gross negligence or beyond, a case can sometimes be made for departing from the default rule, though even here other factors, such as concern to afford relief to those who have already suffered ultimate harms, may counsel caution. In sum, if medical monitoring claims are understood as claims for equitable relief, the courts will be in a better position to integrate policy considerations of the sort identified by Justice Breyer into case-by-case assessments of the appropriateness of class certification, or the granting of relief. By realizing what is happening courts sitting in equity will be in a position to bring a greater degree of coherence and fairness to this burgeoning corner of tort law.

CONCLUSION

We hope that some readers will at least be intrigued by our efforts to rationalize medical monitoring law through a distinction between equitably enforceable duties of conduct and duties of repair enforceable through lump-sum liability awards. We hope the same for our efforts to re-theorize fear-of-disease law in terms of the idea that one must sometimes be vigilant to avoid exposing plaintiffs to threatening situations. Finally, we hope to have helped justify tort law's general refusal to treat increased risk of harm as a

²²³ *Buckley*, 521 U.S. at 442–43; see *id.* at 453 (Ginsburg, J., dissenting).

cognizable injury, and to explain its ban on recovery absent injury in terms of an idea of civil recourse.

At the same time, we worry that some will find these efforts to be old-fashioned, even formalistic. Having created this particular risk, we feel obliged to minimize the chances of its realization. Accordingly, we aim in this Conclusion to temper possible bad impressions by providing a better idea of the more pragmatic side of our perspective.

As noted in the Introduction, one might profitably view American tort law in the first half of the twentieth century as dominated by two figures: Justices Holmes and Cardozo. The two men shared a great deal, both in terms of their careers and, more importantly, their aspirations for the law. Each took his most basic project to be the modernization of law; the readjustment of law to contemporary problems. Each, one would say in modern parlance, was realistic or pragmatic. Still, not all forms of pragmatism are the same and, in fact, the two advocated very different forms of the general disposition toward pragmatism. Holmes's pragmatism combined a respect for clearly stated legal rules as a source of stability, an enthusiasm for social science as a form of knowledge and control, and an understanding of the law purely as an instrument for the realization of desirable social goals.²²⁴ This simultaneous embrace of formalism to guide judicial decision-making, and macro-level, policy-based instrumentalism, helped keep a lid on his deep skepticism about the content of legal concepts and the possibility of any genuine form of knowledge about morality and other normative subjects. Pragmatism as doctrinally constrained instrumentalism promised to address changing conditions and novel problems by keeping the law's goals and its shell of rules relatively continuous, while permitting knowledgeable alteration of hypotheses about which patterns of legal sanctions will most successfully realize those goals. To the extent that this rough sketch captures Holmes, the law and economics approach to torts may fairly claim to be realizing the Holmesian side of pragmatism's legacy for law.

²²⁴ Goldberg, *supra* note 4, at 1432–33, 1458–61 (contrasting Justice Holmes and Justice Cardozo); Goldberg & Zipursky, *supra* note 4, at 1753–56, 1812–24 (contrasting Justice Holmes' and Justice Cardozo's understandings of duty and negligence).

Cardozo shared with Holmes a recognition of the value of stability in law, an abiding interest in social science, and a faith that society could and should be improved in a democratic system through intelligent and self-conscious legal evolution. But Cardozo's body of work represents a quite different form of pragmatism, for he was neither an instrumentalist nor a skeptic. As his occasionally sanctimonious opinions indicate, he took seriously the language of right and duty, and legal concepts more generally. While he shared the pragmatist's faith that all must remain open to revision, this did not cause him to disdain the principles of the common law. Rather it generated a belief that he could work with those principles without being metaphysically locked in. We have referred to this approach as "pragmatic conceptualism" or "conceptualistic pragmatism."²²⁵ It is pragmatic because it commits itself to nothing more than what our legal practices give us; it does not rest upon any claims about "external" legal or moral reality; and it insists on coherence and openness to revision without being doctrinaire or ideological. It is conceptualistic, as opposed to instrumentalist, because it takes legal concepts and legal principles seriously, as fundamental to what makes common law a form of law.

In twentieth-century tort scholarship (and perhaps in legal scholarship generally), it is clear which form of pragmatism won out: Holmes 1, Cardozo 0. Yet even in terms of its own metric of evaluation—usefulness in adjusting the law to change and in solving social problems—we are not persuaded that Holmes deserves the victory. The view of tort law as a means to achieve efficient allocation of resources and activity levels, to serve goals of deterrence and compensation, or to enforce an all-purpose fault principle, each has deep flaws. So pervasive is this problem that huge and important areas of tort—in this instance, those pertaining to risk generation, fear, and medical monitoring—seem largely arbitrary with respect to the instrumentalist understanding of the law. Thus, even an astute and learned judge such as Justice Breyer finds it difficult to articulate a rationale in support of his intuition that Buckley perhaps ought to be able to recover medical monitoring payments from a court-supervised fund.

²²⁵ See discussion *supra* notes 5–6 and accompanying text.

As noted at the outset, even a dyed-in-the-wool Holmesian such as Dean Prosser could, at times, adopt the posture of a pragmatic conceptualist. We have offered, in this paper, another instance of how the pragmatic path not taken—the pragmatic conceptualism of Cardozo—can advance legal analysis and help modernize the law. The topic selected concerns matters outside the purview of courts one hundred or even fifty years ago: the proper response of courts to plaintiffs who claim to have been put at grave risk of injury, in great fear of injury, or to great expense by exposure to hazardous materials. Still, at an interpretive level, it has proved necessary to analyze each of these areas by probing the conceptual structure undergirding the law of torts. With respect to truly inchoate wrongs, we needed to understand the realization requirement, the concept of recourse, and the sense in which tort law conceptually is quite different from a criminal and regulatory law. In the fear-of-injury cases, we needed to resist the temptation to eliminate the distinction between predicate and parasitic injury, and, indeed, to resist the blurring of injury and harm and to ask hard normative questions about the conditions under which emotional distress over the possibility of a future injury or illness is cognizable. We also were required to take seriously the idea of duty—when and how one is obligated to be vigilant of the emotional well-being of others. Finally, on the medical monitoring front, we saw that even to explain the direction of a hard-headed regulatory judge like Justice Breyer in *Buckley*, we needed to address the distinction between equitable relief and damages.

The advantages of this approach go beyond the capacity to interpret doctrine cogently. It provides the possibility of advancing doctrine in a manner that has a better chance of succeeding. If pure risk is the concern, then the appropriate solution is criminal or regulatory law, not tort law. In the fear cases, this means not relying on crude concerns about floodgates and fraud, but thinking seriously about what sorts of emotional injury to guard against, and in what sorts of settings. In the medical monitoring area, it means recognizing that those who create significant risks of injury have a duty to help mitigate those same risks and that those duties might sometimes be appropriately enforced through a court order.

Although the dominant American tradition has been to treat it as an infinitely plastic department of the law, tort law, ironically,

functions less adequately when it is so conceived. Tort law does more good when it progresses with a pragmatic sense of itself as a law of civil recourse. At its best, tort law is progressive yet stable: It articulates and enforces our evolving obligations to one another against the backdrop of an enduring need for a civil forum in which to adjudicate claims and disputes that would otherwise devolve into private vengeance. Armed with this realization, courts and commentators will have a better chance of enabling tort law to realize its function.
