THE YALE LAW JOURNAL

CHRISTOPHER BRETT JAEGER

The Hand Formula's Unequal Inputs

ABSTRACT. Tort cases often hinge on whether the defendant behaved "unreasonably." Tort theorists have long debated what makes behavior unreasonable, with many seeking answers in economic theory or Kantian philosophy. But whether a tort defendant's conduct was unreasonable is typically a question for the jury, and we know little about whether lay jurors' understandings of reasonableness align with tort theorists' beliefs.

This Article addresses this gap in our knowledge, experimentally examining the extent to which economic and Kantian theories capture how laypeople, like jurors, judge reasonableness. The Article explicates and then tests competing economic and Kantian views about the relevance of the various inputs in the well-known Hand Formula ($B < PL$): the burden of taking additional precautions (B), the increased probability of harm occurring in the absence of those precautions (P), and the expected cost to society of the harm should it arise (L). Primarily, it tests the Kantian proposition that the significance of the risks associated with conduct (reflected by the PL side of the Hand Formula) has greater influence on reasonableness judgments than economic theory implies.

Across five experiments, 1,046 participants read simple tort cases that included full information about B, P, and L, in plain language, and then judged whether the defendant acted unreasonably. Participants found defendants' conduct more unreasonable when B was less than PL, as one would expect if participants engaged in some degree of Hand-style, economic balancing of costs and benefits. But, critically, information about P and L had additional, Hand-independent effects on participants' judgments: participants found defendants' conduct more unreasonable as the value of PL—and particularly the value of P—increased, whether B was less than PL or not. A purely economic view of reasonableness cannot explain these findings. The Hand Formula, it seems, is misaligned with how laypeople think about reasonableness. Lay judgments may be influenced by cost justification to a degree, but they are also influenced by Kantian considerations of the significance of risks in ways the Hand Formula does not capture.

In short, the Hand Formula's inputs do not have equal weight in lay decision-makers' minds. P is most influential, though decision-makers may not be conscious of its outsized influence. These findings have theoretical and practical implications, including informing the ongoing normative debate as to how tort law should understand reasonableness, underscoring the importance of juries' democratic role in negligence litigation, and highlighting types of evidence and arguments that may be particularly likely to persuade jurors.

AUTHOR. Associate Professor of Law, Baylor Law School; J.D., Vanderbilt Law School; Ph.D. (Psychology), Vanderbilt University. I am grateful to Jessica Asbridge, Gilat Bachar, Edith

Beerdsen, Joseph Blocher, Lauren Clatch, Neal Feigenson, Steve Gilles, Tonja Jacobi, Ted Jaeger, Rachel Kincaid, Heidi Liu, Jamie Macleod, Luke Meier, Jeff Rachlinski, Jennifer Robbennolt, Matthew A. Shapiro, Jay Silver, Ken Simons, Rod Sturdivant, Stephanie Tang, Jim Underwood, Richard Wright, and to participants in the Lay Understanding of Justice Concepts Session at Law and Society 2024, the Torts and Civil Procedure Session at the 2024 Southeastern Association of Law Schools Annual Meeting, the 2024 Association of American Law Schools (AALS) Jurisprudence Section's Junior Works-in-Progress Workshop, the 2024 Central States Law Schools Association Conference, the 2024 Conference on Empirical Legal Studies, and colloquia at Baylor Law School and Emory Law School for helpful comments and conversations about this project. I am also grateful to Sammy Lutes, Jack Freese, and Justin Garber for excellent research assistance and to the editorial staff of the *Yale Law Journal* for exceptional comments and editorial work. I did not use any form of artificial intelligence in creating this work. This Article was named cowinner of the 2025 AALS Scholarly Papers Competition.

ARTICLE CONTENTS

INTRODUCTION

Legal disputes, and tort negligence disputes in particular, often hinge on whether a defendant's actions were "unreasonable." But what does it mean for actions to be "unreasonable"? What criteria do legal decision-makers, often lay jurors, use to classify actions as reasonable or not?

Legal theorists have debated the criteria of reasonableness for decades,[1] with the most active fault line in the debate situated between economic perspectives (grounded in utilitarianism) and Kantian perspectives (grounded in deontological ethics).[2] Much of this debate has been normative, focused on how law should define reasonableness. But the debate has also been descriptive.[3] Both economic and Kantian views aim to provide practical interpretations of tort law. Proponents of each perspective routinely cite specific cases that allegedly have been decided consistently with their preferred view.[4] Proponents of each contend that fundamental features of negligence law demonstrate that it is animated by their preferred view.[5] And proponents of each have argued that their view is superior

1. For scholarly discussion of the criteria for reasonableness, see generally, for example, Alan D. Miller & Ronen Perry, *The Reasonable Person*, 87 N.Y.U. L. Rev. 323 (2012); Benjamin C. Zipursky, *Reasonableness in and out of Negligence Law*, 163 U. Pa. L. Rev. 2131 (2015); Kevin P. Tobia, *How People Judge What Is Reasonable*, 70 Ala. L. Rev. 293 (2018); Kenneth W. Simons, *The Hegemony of the Reasonable Person in Anglo-American Tort Law*, *in* 1 Oxford Studies in Private Law Theory 45 (Paul B. Miller & John Oberdiek eds., 2020); and Mark D. Alicke & Stephanie H. Weigel, *The Reasonable Person Standard: Psychological and Legal Perspectives*, 17 Ann. Rev. L. & Soc. Sci. 123 (2021).

2. *See, e.g.*, Cristina Carmody Tilley, *Tort Law Inside Out*, 126 Yale L.J. 1320, 1324-25 (2017) ("Broadly speaking, tort theory has split into two camps: economists view tort as a method of encouraging efficient private behavior, while philosophers and political scientists view it as a method of achieving a kind of moral justice."); Miller & Perry, *supra* note 1, at 350 ("The Kantian definition of reasonableness is the strongest rival of the economic definition.").

3. *See* Patrick J. Kelley, *Who Decides? Community Safety Conventions at the Heart of Tort Liability*, 38 Clev. St. L. Rev. 315, 319-20 (1990) (reviewing prominent tort theories, including Kantian and economic theories, and noting that "[e]ach theorist claimed that his theory was descriptive").

4. *See, e.g.*, William M. Landes & Richard A. Posner, *The Positive Economic Theory of Tort Law*, 15 Ga. L. Rev. 851, 884-85, 892-903 (1981) (reviewing cases in support of the argument that the Hand Formula is descriptive); Stephen G. Gilles, *The Invisible Hand Formula*, 80 Va. L. Rev. 1015, 1016 n.4 (1994) (collecting judicial opinions endorsing the Hand Formula or similar formulations of negligence); Richard W. Wright, *Hand, Posner, and the Myth of the 'Hand Formula,'* 4 Theoretical Inquiries L. 145, 146-48, 211-23 (2003) (arguing that case law is better explained by Kantian "equal freedom" than by economic efficiency).

5. *See, e.g.*, Richard A. Posner, *A Theory of Negligence*, 1 J. Legal Stud. 29, 52 (1972) (contending that "the basic formal structure of the negligence system broadly supports an economic theory of negligence"); Richard W. Wright, *The Standards of Care in Negligence Law*, *in*

because it is more intuitive and better aligns with how the relevant legal decision-makers—typically lay jurors—understand reasonableness.[6] Indeed, the descriptive debate in many ways hinges on the open empirical question of how tort jurors apply the reasonableness standard.[7]

This Article addresses that open question, making both empirical and theoretical contributions. Empirically, this Article reports the first experiments investigating whether economic or Kantian theory better reflects how laypeople judge reasonableness. In laying the groundwork for the experiments, this Article carefully engages with both economic and Kantian theories to derive testable

---

PHILOSOPHICAL FOUNDATIONS OF TORT LAW 249, 256 (David G. Owen ed., 1995) ("It is the equal freedom conception of reasonableness . . . that is reflected in actual tort-law doctrines and decisions."); Steven Hetcher, *Non-Utilitarian Negligence Norms and the Reasonable Person Standard*, 54 VAND. L. REV. 863, 867 (2001) (noting broadly that advocates of Kantian and economic views have "each provide[d] a positive account of tort law that, as it happens, precisely dovetails with [their] prescribed normative account").

6. *See, e.g.*, Gilles, *supra* note 4, at 1020 ("Some proponents of Hand Formula negligence believe that jurors, if given only a reasonable person instruction, intuitively will arrive at results consistent with cost-benefit analysis at least as often as they would if given an explicit Hand Formula instruction. Proponents of 'community values' conceptions of negligence claim, to the contrary, that jurors will look to existing social practices and conventions to fill in the reasonable person standard."); Brian Sheppard, *The Reasonableness Machine*, 62 B.C. L. REV. 2259, 2296 (2021) (citing WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 85-86 (1987)) (noting some scholars "have maintained that the Hand Test is the very best description of the [reasonable-person standard] in torts doctrine"); Arthur Ripstein, *Reasonable Persons in Private Law*, *in* REASONABLENESS AND LAW 255, 273 (Giorgio Bongiovanni, Giovanni Sartor & Valentini Chiara eds., 1st ed. 2009) ("A reasonable person[] would not think in the terms suggested by the Learned Hand test . . . [but] takes such care as is required by a like security and liberty for all."); Richard W. Wright, *Justice and Reasonable Care in Negligence Law*, 47 AM. J. JURIS. 143, 187-88 (2002) ("[I]t clearly is the disagreement of judges and jurors with the Hand-formula definition of reasonableness that explains both judges' refusal to give Hand-formula instructions to jurors and the well-grounded expectation that jurors would often refuse to follow such instructions if they were given."). While some scholars in both the economic and Kantian camps focus more on judges' decisions than on juries', reasonable care generally remains the province of juries. *See* RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM § 8(b) (A.L.I. 2010) ("When, in light of all the facts relating to the actor's conduct, reasonable minds can differ as to whether the conduct lacks reasonable care, it is the function of the jury to make that determination."). And even where judges take the reasonable-care issue from the jury, lay understanding likely highlights at least some conceptual underpinnings of judicial approaches. *See* Kevin P. Tobia, *Law and the Cognitive Science of Ordinary Concepts*, *in* LAW AND MIND: A SURVEY OF LAW AND THE COGNITIVE SCIENCES 86, 86 (2021) ("Laypeople's common-sense understandings . . . are at the root of *many* important legal concepts . . . including . . . reasonableness . . . .").

7. Gilles, *supra* note 4, at 1020 ("At one level, the debate over the positive meaning of negligence turns on an empirical question: confronted with a blank reasonable person instruction, what will jurors do?").

predictions as to the types of information that are relevant when judging reasonableness.

Economic theorists have generally identified their reasonableness criteria more concretely than Kantians. On the leading economic view, the reasonableness of a behavior depends on aggregate cost-benefit analysis.[8] People act unreasonably when their actions create risks that are not cost-justified. In the context of tort negligence cases, cost-benefit analysis is typically applied to the defendant's failure to take some specific precaution.[9] The failure to take the precaution is "unreasonable"—and the defendant may therefore be found liable for negligence—if "the reduction in expected accident costs [due to the precaution] would have been greater than the costs of the precaution."[10] This analysis is expressed through the well-known Hand Formula, which instructs that the failure to take a particular precaution is unreasonable when the burden of taking the precaution (B) is less than the increased probability of harm occurring in the absence of the precaution (P) multiplied by the expected cost to society of the harm should it arise (L)—that is, when $B < PL$—but reasonable where the burden is greater than that product—when $B > PL$.[11]

The most prominent alternatives to this economic account of reasonableness are Kantian.[12] Kantian accounts focus on principles of justice rather than cost-

8. *See, e.g.*, Posner, *supra* note 5, at 32-33 (framing Hand balancing in terms of "overall economic value" and what makes "society" better off); Kenneth W. Simons, *Tort Negligence, Cost-Benefit Analysis, and Tradeoffs: A Closer Look at the Controversy*, 41 LOY. L.A. L. REV. 1171, 1172 (2008) ("For most utilitarians and adherents to law and economics . . . to say that people should not be negligent is to say that they should minimize the aggregate expected costs of their activities—specifically, they should minimize the sum of the costs of accidents and the costs of preventing accidents."). For an alternative, nonaggregative approach to the Hand Formula, see generally Emad H. Atiq, *The Disaggregated Hand Formula*, 114 CALIF. L. REV. (forthcoming 2026), https://ssrn.com/abstract=5152136 [https://perma.cc/Z9PJ-8PMY].

9. *See* Mark F. Grady, *Untaken Precautions*, 18 J. LEGAL STUD. 139, 139 (1989) ("The key question that courts ask is what particular precautions the defendant could have taken but did not."); Simons, *supra* note 8, at 1172-73 (noting the economic view that people should take a precaution only if the marginal costs of that precaution (in terms of tangible costs or lost benefits) are less than the marginal benefits (in terms of reduced risk of injury)). It is not clear that every negligence claim is premised on the failure to take a precaution. For more on this point, see generally Christopher Brett Jaeger, *A Two-System Theory of Negligence*, 93 U. CHI. L. REV. (forthcoming 2026) (on file with author). This Article focuses on the large set of negligence cases that do.

10. Stephen G. Gilles, *On Determining Negligence: Hand Formula Balancing, the Reasonable Person Standard, and the Jury*, 54 VAND. L. REV. 813, 818 (2001).

11. *See* United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947) (providing Judge Learned Hand's most famous articulation of the Hand Formula).

12. There are, of course, many other accounts of reasonableness. Both social contractarianism and reciprocity-based views share some conceptual overlap with Kantian theory. *See, e.g.*, Gregory

benefit analysis.[13] They proceed from the idea that all human beings are free and equal, to be treated as ends in themselves rather than means to others' ends.[14] Under Kantian accounts, an actor behaves unreasonably whenever the actor encroaches on others' equal freedom by imposing "significant" (or "substantial") risks on others without their consent.[15]

But how do Kantian accounts distinguish significant (and thus potentially actionable) risks from insignificant ones? Think of this issue in terms of the Hand Formula inputs of *B*, *P*, and *L*. While Kantian views are not generally concerned with Hand-style cost-benefit analysis, they are concerned with whether an actor imposes significant risks on others without consent.[16] To impose risk on others is to impose some increased probability of harm on them—*P*. The Kantian concern with "significant" risks may imply that the severity of the harm being risked—*L*—is also relevant in evaluating whether the conduct that created the risk was reasonable.

The burden of taking a precaution, *B*, matters less to Kantians. On the strongest Kantian views, the burden is irrelevant: if the risks imposed by a defendant's conduct are "significant" (based on a combination of *P* and *L*), it makes no difference how costly it would be for the defendant to mitigate them.[17] Under

C. Keating, *Reasonableness and Rationality in Negligence Theory*, 48 STAN. L. REV. 311, 359-68 (1996) (discussing social-contract theory); T.M. SCANLON, WHAT WE OWE TO EACH OTHER 191-97 (1998) (same); George P. Fletcher, *Fairness and Utility in Tort Theory*, 85 HARV. L. REV. 537, 569-73 (1972) (discussing reciprocity). For sources reviewing various perspectives on reasonableness, see *supra* note 1.

13. *See* Ernest J. Weinrib, *Deterrence and Corrective Justice*, 50 UCLA L. REV. 621, 633 (2002) ("Kant's treatment of right is consonant with—and is in part a philosophical elucidation of—corrective justice.").

14. *See, e.g.*, Ernest J. Weinrib, *Toward a Moral Theory of Negligence Law*, 2 LAW & PHIL. 37, 54 (1983) ("[O]n the Kantian interpretation the defendant's failure to exercise due care is a failure to give equal consideration to the plaintiff and is thus a wrong directed against him."); Wright, *supra* note 6, at 164 (working from the Kantian imperative to treat humanity "always as an end and never as a means only" (quoting IMMANUEL KANT, FOUNDATIONS OF THE METAPHYSICS OF MORALS 47 (Lewis White Beck trans., Bobbs-Merrill Co. 1959) (1785)).

15. *See, e.g.*, Wright, *supra* note 5, at 256 ("[I]t is impermissible to use someone as a mere means to your ends by exposing him (or his resources) to significant foreseeable unaccepted risks.").

16. *Id.*

17. *See, e.g.*, Ernest J. Weinrib, *Correlativity, Personality, and the Emerging Consensus on Corrective Justice*, 2 THEORETICAL INQUIRIES L. 107, 116 (2001) ("[T]he unreasonableness of the risk created by the defendant is seen in terms of the probability and the gravity of its effects on others . . . .") [hereinafter Weinrib, *Correlativity, Personality, and the Emerging Consensus on Corrective Justice*]. Weinrib argues that English courts generally and correctly ignore *B* and "focus[] narrowly on the risk, consisting in the combination of P and L." ERNEST J. WEINRIB, THE IDEA OF PRIVATE LAW 148 (1995) [hereinafter WEINRIB, THE IDEA OF PRIVATE LAW]. To the extent English courts allow some role for *B* in cases where risks are not fantastical but

any circumstances, imposing a significant, unaccepted risk on others encroaches on their equal freedom and is therefore unreasonable. Softer Kantian views, which allow some role for the extent of the burden, still assign *B* less importance than *P* and *L*.[18] In sum, on Kantian views, a subset of Hand Formula inputs are particularly relevant to evaluating reasonableness: *P* and *L* (perhaps especially *P*) play a larger role in the calculus of negligence than the Hand Formula implies.[19]

Because the question of whether a negligence defendant acted reasonably is typically one for the jury, assessing whether the economic or Kantian view better describes the law of negligence turns in large part on lay jurors' conceptions of reasonableness. The Hand Formula predicts that lay jurors weigh (or act as though they weigh) the burdens of taking precautions (*B*) against the risks of proceeding without them (*PL*); Kantian views predict they focus more, or perhaps exclusively, on the risks (*PL*). This Article tests these predictions, using experimental methods to probe the criteria lay decision-makers use to evaluate reasonableness.

Across five experiments, 1,046 participants assessed negligence cases that included full information about *B*, *P*, and *L* in plain language. Participants judged defendants more negligent in situations where *B* was less than *PL* than in situations where *B* was greater than *PL*, as one would expect if participants engaged in some degree of Hand-style cost-benefit analysis. But the Hand Formula did not capture the full effect of *B*, *P*, and *L*. Information about *P* and *L*—particularly *P*—influenced participants' judgments above and beyond the weight accorded by the Hand Formula. This Kantian effect appeared whether the case involved risks of property damage or physical injury, whether those risks were concentrated on one identified person or spread across multiple unidentified people, and whether the defendant was an individual or a corporate entity. And it persisted even when participants were told exactly what cost-benefit analysis indicated the defendant should have done. In sum, my studies provide experimental evidence that *B*, *P*, and *L* do not factor into the calculus of negligence quite as the Hand Formula assumes: *PL*, and particularly *P*, plays an outsized, Kantian role.

---

remain too small to be considered significant, see Weinrib, *Correlativity, Personality, and the Emerging Consensus on Corrective Justice*, *supra*, at 116. Weinrib considers this a mistake. WEINRIB, THE IDEA OF PRIVATE LAW, *supra*, at 150 n.9 (noting that "even the modest role the English approach assigns to Learned Hand's B may be excessive"). *See also* Stephen G. Gilles, *The Emergence of Cost-Benefit Balancing in English Negligence Law*, 77 CHI.-KENT L. REV. 489, 496 (2002) (describing the Kantian view that actors are negligent when they impose a substantial, foreseeable, unaccepted risk on others, "without regard to how difficult it would have been to guard against that risk").

18. *See infra* notes 86-87 and accompanying text.

19. *See infra* Section I.B.2.

In addition to informing important debates within the legal academy, these findings also speak to practicing tort lawyers. While more work is needed to build the bridge from vignette studies to legal practice, these findings suggest that lawyers trying tort cases might strategically emphasize evidence about the probability of harm over evidence about the severity of harm or costs of additional precautions. Further, a more refined sense of lay jurors' concept of reasonableness can help lawyers better value cases for settlement.

The Article proceeds as follows. Part I discusses the role of the reasonableness standard in tort law. It explores economic and Kantian views of that standard, identifying what each view predicts about the role of *B*, *P*, and *L* information in reasonableness judgments. Part II reports five original experiments testing whether *B*, *P*, and *L* information influences lay reasonableness judgments in the ways that economic and Kantian views predict. Part III summarizes key experimental findings and discusses their implications, both for the scholarly debate about reasonableness and for the everyday practice of tort law. Part III also discusses promising directions for future research. Part IV concludes.

I. REASONABLENESS AND UNREASONABLENESS IN TORT LAW

This Part discusses reasonableness in tort practice, in legal theory, and in empirical legal research. Section I.A explains why tort liability often hinges on lay judgments about whether conduct is reasonable. Section I.B proceeds to describe economic and Kantian views of what it means to behave reasonably, identifying articles and cases that have endorsed and applied each view. Finally, Section I.C reviews empirical research on how laypeople understand and judge reasonableness, setting the stage for this Article to investigate empirically the degree to which lay reasonableness judgments align with economic or Kantian views.

A. The Role of Reasonableness in Tort Practice

In the American legal system, the default is that people who get injured bear the cost. Harms "lie where they fall."[20] Tort claims are exceptions to this default. When a plaintiff prevails on a tort claim, the court will order the defendant to pay at least some of the costs associated with the plaintiff's injuries.[21]

The most common tort claim is negligence. To prevail on a negligence claim, a plaintiff must generally demonstrate four elements: (1) the plaintiff suffered

20. Oliver Wendell Holmes, Jr., The Common Law 144 (Bos., Little, Brown & Co. 1881).

21. *See* John C.P. Goldberg, Anthony J. Sebok & Benjamin C. Zipursky, Tort Law: Responsibilities and Redress 3 (2004).

an injury; (2) the defendant owed the plaintiff a duty to act with care; (3) the defendant breached that duty; and (4) the breach caused the plaintiff's injury.[22]

Reasonableness plays a critical role in analyzing the duty and breach elements. The duty element is typically resolved by the court, with the general rule being that "an actor whose conduct creates risks of physical harm to others has a duty to exercise reasonable care."[23] Defendants breach their duty when they fail to conduct themselves with reasonable care—that is, when they act unreasonably.[24]

The breach element is customarily treated as an issue for the jury, meaning that it is typically lay jurors who categorize a negligence defendant's conduct as "reasonable" or "unreasonable."[25] Courts have justified this custom by observing that determining whether a defendant's conduct was reasonable often requires experience-driven social judgments best made by jurors from the relevant community.[26] Jurors typically receive minimal and open-ended instruction on this task.[27] Consequently, "the negligence standard of reasonable care is essentially defined by the lay understanding of jurors"—jurors' beliefs about reasonableness

---

22. *See* John C.P. Goldberg & Benjamin C. Zipursky, *The* Restatement (Third) *and the Place of Duty in Negligence Law*, 54 VAND. L. REV. 657, 658 (2001).

23. RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM § 6 cmt. b (A.L.I. 2010).

24. For the sake of concise expression, scholars often equate "unreasonable" behavior with behavior lacking reasonable care. *See, e.g.*, Teneille R. Brown, *Minding Accidents*, 94 U. COLO. L. REV. 89, 101 (2023) (describing a cluster of terms frequently employed interchangeably in negligence cases). It is, of course, possible for an actor to behave "unreasonably" in ways unrelated to carefulness, but in this Article, "unreasonable" behavior means behavior that falls short of the standard of reasonable care. The breach analysis focuses on the defendant's conduct, and not on the defendant's mental state. Timothy D. Lytton, *Rules and Relationships: The Varieties of Wrongdoing in Tort Law*, 28 SETON HALL L. REV. 359, 366 ("Establishing negligence . . . does not require any evidence or proof of defendant's state of mind.").

25. RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM § 8(b) (A.L.I. 2010).

26. *See id.* § 8 cmt. b ("The jury is assigned the responsibility of rendering such judgments partly because several minds are better than one, and also because of the desirability of taking advantage of the insight and values of the community, as embodied in the jury, rather than relying on the professional knowledge of the judge."); *see also infra* note 296 and accompanying text (discussing democratic deference to community mores).

27. *See* Ashley M. Votruba, *Will the Real Reasonable Person Please Stand Up?: Using Psychology to Better Understand How Juries Interpret and Apply the Reasonable Person Standard*, 45 ARIZ. ST. L.J. 703, 705 (2013) (explaining that jurors "are largely left to their own devices to decide what is considered negligent behavior . . . with only the vague, undefined concept of the reasonable person as their guide"); Gilles, *supra* note 4, at 1019 ("[T]he reasonable person standard is given to the jury without elaboration."). For a survey of negligence jury instructions, see generally Patrick J. Kelley & Laurel A. Wendt, *What Judges Tell Juries About Negligence: A Review of Pattern Jury Instructions*, 77 CHI.-KENT L. REV. 587 (2002).

define the boundaries of the law.[28] If the jury finds that the defendant's behavior falls short of reasonable care, the defendant will be liable for negligence if the other elements of the claim (injury, duty, and causation) are also satisfied. If the jury deems the defendant's behavior reasonable, however, the defendant cannot be held liable for negligence.

B. Dueling Theories of Reasonableness

Legal theorists have long debated how tort law's reasonableness standard is and should be understood.[29] This Article homes in on two prominent perspectives within tort scholarship: economic and Kantian.[30] This Section reviews scholarship and case law related to these views and identifies what each implies about the relevant criteria for judging reasonableness.

1. Economic Reasonableness (Does PL *Exceed* B*?)*

Economic views are built on the idea that conduct's reasonableness depends on aggregate cost-benefit analysis. This idea's intellectual origins lie in act utilitarianism, or the philosophy that the moral act is the one that "among alternatives open to [the] agent . . . has the best consequences."[31] In a 1915 article, Henry T. Terry suggested that the reasonableness of conduct could be evaluated by using factors to balance the "foreseeable risks created by the conduct . . . against the conduct's expected utility."[32] This type of cost-benefit balancing was also reflected in the *Restatement (First) of Torts*,[33] before it was

28. Mark A. Geistfeld, *Folk Tort Law*, *in* RESEARCH HANDBOOK ON PRIVATE LAW THEORY 338, 338 (Hanoch Dagan & Benjamin C. Zipursky eds., 2020); *see also* Kenneth S. Abraham, *The Trouble with Negligence*, 54 VAND. L. REV. 1187, 1191 (2001) (describing the negligence decision "in an unbounded tort case" as typically "an act of discretionary norm creation by the finder of fact").

29. A comprehensive review of this literature is beyond the scope of this Article. For sources reviewing the literature, see *supra* note 1.

30. *See, e.g.*, Stephen R. Perry, *The Moral Foundations of Tort Law*, 77 IOWA L. REV. 449, 449 (1992) (describing two main categories of tort theories, one viewing tort liability as a means for attaining public-policy goals such as economic efficiency and the other viewing tort liability as "based on individual moral rights").

31. Richard F. Bernstein, *Legal Utilitarianism*, 89 ETHICS 127, 127 (1979).

32. Henry T. Terry, *Negligence*, 29 HARV. L. REV. 40, 42-44 (1915); *cf.* Heidi Li Feldman, *Prudence, Benevolence, and Negligence: Virtue Ethics and Tort Law*, 74 CHI.-KENT L. REV. 1431, 1442 (2000) (arguing for an alternative interpretation of Terry's framework).

33. *See* Michael D. Green, *Negligence = Economic Efficiency: Doubts* >, 75 TEX. L. REV. 1605, 1606, 1622-29 (1997) (discussing cost-benefit balancing in the First Restatement).

famously distilled into the three-factor Hand Formula in *United States v. Carroll Towing Co.*[34]

Carroll Towing involved the sinking of a barge off a pier in New York Harbor.[35] The case was complex, involving alleged missteps by numerous actors, but for present purposes, we can focus on one issue: whether the owner of the barge was at fault for failing to post a bargee on board when it sank. If on board, a bargee would have called for help from other nearby boats, which could have safely beached the barge and prevented many of the ensuing losses.[36] With precedent divided as to whether leaving a barge unmanned breached a barge owner's duty of care, Judge Learned Hand sought to "derive a general liability test" to address the issue.[37] He explained that the "owner's duty, as in other similar situations, to provide against resulting injuries is a function of three variables: (1) The probability that [the barge] will break away [P]; (2) the gravity of the resulting injury, if she does [L]; (3) the burden of adequate precautions [B]."[38] Judge Hand went on to state that "liability depends upon whether B is less than L multiplied by P"—whether the burden of adequate precautions was less than the expected costs of proceeding without those precautions.[39] In short, "reasonable care requires taking precautions if and only if those precautions are cost-justified," meaning $B < PL$.[40] With the boat in a "crowded harbor where moored barges are constantly being shifted about," the court concluded that the high probability and gravity of a barge breaking free made it a "fair requirement that [the barge owner] should have a bargee aboard (unless he had some excuse for his absence), during the working hours of daylight."[41]

The idea that conduct is reasonable when cost-justified and unreasonable when not cost-justified was quickly embraced in the legal academy[42] and

34. 159 F.2d 169 (2d Cir. 1947); *see also* Conway v. O'Brien, 111 F.2d 611, 611-13 (2d Cir. 1940) (articulating the factors involved in the Hand Formula six years prior to their more famous articulation in *Carroll Towing*); Patrick J. Kelley, *The* Carroll Towing Company *Case and the Teaching of Tort Law*, 45 ST. LOUIS U. L.J. 731, 744 (2001) ("Judge Hand's statements . . . of the factors relevant to a negligence determination could be seen as an elegant[], concise summary of the [First] Restatement's more elaborate cost-benefit standard . . . .").

35. 159 F.2d at 169, 170-71, 173.

36. *Id.* at 172.

37. Kelley, *supra* note 34, at 739.

38. *Carroll Towing*, 159 F.2d at 173.

39. *Id.*

40. Zipursky, *supra* note 1, at 2151.

41. *Carroll Towing*, 159 F.2d at 173-74.

42. *See* Benjamin C. Zipursky, *Sleight of Hand*, 48 WM. & MARY L. REV. 1999, 2001-04 (2007) (describing the economic theory of torts as "the most celebrated within the legal academy");

"canonized in the law-and-economics literature."[43] The Hand Formula places untaken precautions like manning the barge at the center of negligence analysis,[44] suggesting that a negligence plaintiff's first move should be identifying some untaken precaution that "would have significantly reduced the probability of the accident at low cost."[45] If an untaken precaution would have prevented accidents like the one that injured the plaintiff,[46] and the cost of that untaken precaution was less than the corresponding reduction in expected losses from those accidents, the failure to take the precaution is unreasonable.[47]

For the Hand Formula to generate a determinate answer on the issue of reasonableness, the burdens of precautions (B) and the expected gravity of harm should harm occur (L) need to be quantified on a common scale. The preferred common scale is dollars.[48]

To be clear, proponents contend that the Hand Formula can and does usefully guide decision-makers even when the precise values of B, P, and L are undefined (in dollars or otherwise). They posit that, in the absence of precise values, decision-makers nevertheless intuitively weigh whatever information they have about costs and benefits—roughly balancing, for example, their sense that additional precaution would be moderately costly against a sense that an accident

Kelley, *supra* note 34, at 757 (noting the prominent place of *Carroll Towing* in American tort classes and casebooks).

43. Allan M. Feldman & Jeonghyun Kim, *The Hand Rule and* United States v. Carroll Towing Co. *Reconsidered*, 7 AM. L. & ECON. REV. 523, 523 (2005).

44. "The key question that courts ask [in negligence cases] is what particular precautions the defendant could have taken but did not." Grady, *supra* note 9, at 139. "[E]ach element of a plaintiff's specific negligence case depends on how his or her lawyer has specified the untaken precaution." *Id.* at 141. The burden (B) required to take the untaken precaution, which includes both tangible costs and foregone benefits, is weighed against the corresponding change in risk (PL). Mark F. Grady, *A New Positive Economic Theory of Negligence*, 92 YALE L.J. 799, 801-03 (1983).

45. John Prather Brown, *Toward an Economic Theory of Liability*, 2 J. LEGAL STUD. 323, 334-35 (1973).

46. The Hand Formula can be applied to evaluate the reasonableness of any potential precaution, even precautions that would not likely have prevented the plaintiff's injury. Of course, if a precaution would not have prevented the plaintiff's injury, the plaintiff's claim should fail on *causation* grounds. Grady, *supra* note 44, at 827-28, 828 n.67.

47. *See, e.g.*, Feldman & Kim, *supra* note 43, at 523.

48. *See, e.g.*, Zipursky, *supra* note 1, at 2159 ("On the economic view, the aggregative view is specified further by reference to social wealth as the metric."). There are other, more abstract units one could, in theory, use to quantify burdens and gravity, such as utility or social welfare. Gilles, *supra* note 10, at 818-19. As a practical matter, though, such alternatives are often ultimately operationalized in dollars. Michael Pressman, *The Compatibility of Forward-Looking and Backward-Looing Accounts of Tort Law*, 15 U.N.H. L. REV. 45, 53 (2016).

was very unlikely to happen but would be severe if it did.[49] While the application of the Hand Formula fails to produce an unambiguous result in such situations, proponents urge that it nevertheless captures something about how people judge reasonableness.[50] And in the rare cases when *B*, *P*, and *L* are clearly defined, proponents contend that the Hand Formula is not only normatively desirable, but that it also provides a descriptively sound account of how people judge reasonableness.[51]

It merits reemphasizing that the Hand Formula considers costs and benefits in the aggregate.[52] As theorist Richard W. Wright observed, "[C]onduct is efficient (hence reasonable and not negligent) if it maximizes the total sum of expected benefits minus costs; it does not matter who bears the costs and who gains the benefits."[53] It likewise does not matter how concentrated or diffuse

---

49. McCarty v. Pheasant Run, Inc., 826 F.2d 1554, 1557 (7th Cir. 1987) (explaining that because lawyers do not typically present Hand quantities, "[f]or many years to come juries may be forced to make rough judgments of reasonableness, intuiting rather than measuring the factors in the Hand Formula"); Kenneth W. Simons, *The Hand Formula in the Draft* Restatement (Third) of Torts*: Encompassing Fairness as Well as Efficiency Values*, 54 VAND. L. REV. 901, 908 (2001) (contending that the Hand Formula need not be understood in reductionist terms to mean that actors "literally 'calculate' the relevant costs and benefits").

50. *See, e.g.*, Gabriel Weil, *Efficiency, Fairness, and the Externalization of Reasonable Risks: The Problem with the Learned Hand Formula*, 75 S.C. L. REV. 155, 157 (2023) ("Even in the majority of cases where quantification of the B, P, and L parameters appears infeasible then, many law students, lawyers, and judges accept that some loose approximation of the Hand formula is the ideal that the law is and should be striving for.").

51. *McCarty*, 826 F.2d at 1557 (asserting that the Hand Formula tends to have "greater analytic than operational significance" only because ordinarily, "the parties do not give the jury the information required to quantify the variables that the Hand Formula picks out as relevant"); Gilles, *supra* note 10, at 853 ("On the Posnerian view, full information about the Hand Factors would yield 'certain results' because that information would consist of monetized values . . . leaving no room for judgment.").

52. Miller & Perry, *supra* note 1, at 330-31 ("In fact, the Hand formula is aggregative: It balances social advantages against social disadvantages."). There has been scholarly discussion as to how broadly various Hand Formula inputs should be construed. For instance, scholars have debated whether the likely injuries in the event of an accident—*L*, in Hand Formula terms—include expected losses *to the tortfeasor* or only expected losses *to others*. *See, e.g.*, Gary T. Schwartz, *Contributory and Comparative Negligence: A Reappraisal*, 87 YALE L.J. 697, 723-24 (1979); Robert Cooter & Ariel Porat, *Does Risk to Oneself Increase the Care Owed to Others?: Law and Economics in Conflict*, 29 J. LEGAL STUD. 19, 28 (2000); Kenneth W. Simons, *The Puzzling Doctrine of Contributory Negligence*, 16 CARDOZO L. REV. 1693, 1725-28 (1995). Scholars have also disagreed as to whether *B* includes only the cost to the defendant of increased precaution or the social cost. *See* Ronen Perry, *Harmful Precautions*, 99 NOTRE DAME L. REV. 153, 155, 167-69 (2023). I designed the vignettes in my studies to avoid implicating these issues, though future empirical work could examine them.

53. Wright, *supra* note 5, at 251; *see also* Grady, *supra* note 9, at 139 (arguing that under the Hand Formula, "the critical inquiry is whether the defendant has chosen a level of precaution that globally minimizes social cost").

those benefits and costs are. On the economic view, then, I should not spend my $500 on precautions that would prevent an expected loss of $499 to you, nor should I spend my $500 on precautions that would prevent an expected loss of $1 for each of 499 individuals.

Rhode Island Hospital Trust National Bank v. Zapata Corp.[54] illustrates the Hand Formula in action. The plaintiff, Zapata Corporation, suffered losses due to forged checks drawn on its account with the defendant bank.[55] Zapata sought to recover on the theory that the bank's system for detecting forgery — examining signatures for only a small percentage of checks chosen at random — was inadequate.[56] Zapata contended that, to satisfy the standard of reasonable care, the bank needed to have taken an additional precaution: individual examination of each check drawn.[57] In affirming judgment for the bank, the First Circuit noted that individual examination of each check would have cost the bank approximately $125,000 more per year than its random screening approach (*B* of $125,000), with little, if any, corresponding reduction in the erroneous payment of forged checks (*PL* close to $0).[58] Zapata argued the costs the bank saved were irrelevant — that the Hand Formula's *B* input should have no bearing when assessing whether the bank took reasonable care.[59] The Court disagreed, reasoning that "what is reasonable or unreasonable . . . is often a matter of costs of prevention compared with correlative risks of loss."[60]

The American Law Institute (ALI) largely adopted the Hand Formula as the standard for evaluating (un)reasonableness in the *Restatement (Third) of Torts: Liability for Physical and Emotional Harm*, published in 2010:

> A person acts negligently if the person does not exercise reasonable care under all the circumstances. Primary factors to consider in ascertaining whether the person's conduct lacks reasonable care are the foreseeable likelihood that the person's conduct will result in harm, the foreseeable severity of any harm that may ensue, and the burden of precautions to eliminate or reduce the risk of harm.[61]

---

54. 848 F.2d 291 (1st Cir. 1988).

55. *Id.* at 292.

56. *Id.* at 293-94.

57. *Id.* at 295.

58. *Id.* at 294.

59. *Id.* at 295.

60. *Id.*

61. Restatement (Third) of Torts: Liab. for Physical & Emotional Harm § 3 (A.L.I. 2010). Though debated, some scholars have argued that the *Restatement (Third)* adopts an

However, both the Hand Formula and its adoption by the ALI have drawn significant criticism. Many have observed that it is the rare case where *B*, *P*, and *L* can be quantified well enough to make the formula practically useful.[62] Some have argued the Hand Formula fails to properly account for precautions that other litigants might have taken, or that litigants might have taken together.[63] But likely the loudest critique is that adopting the Hand Formula put the Third Restatement far out of step with real-world jurisprudence by "dramatically overstat[ing] the role of . . . cost-benefit analysis in the reasonable person standard."[64] As Steven Hetcher observed, the Hand Formula "appears to have played almost no role in the determinations of the countless jury decisions over the years that, taken together, constitute the substance of the standard for breach in American tort law."[65]

Nevertheless, the Hand Formula remains tremendously influential within the academy and particularly in the first-year classroom.[66] Wright has suggested that most legal scholars assume, at least implicitly, that the Hand Formula "is descriptively correct."[67] Proponents of the Hand Formula argue that even if jurors are not instructed to apply it, they do so intuitively and implicitly.[68]

"unconstrained, reductionist cost-benefit test of reasonableness in negligence law." Richard W. Wright, *Negligence in the Courts: Introduction and Commentary*, 77 CHI.-KENT L. REV. 425, 430-31 (2001).

62. Gilles, *supra* note 10, at 818 ("As Learned Hand pointed out, it will often be impossible to arrive at a 'quantitative estimate' of any of the factors comprising the Hand Formula." (quoting Moisan v. Loftus, 178 F.2d 148, 149 (2d Cir. 1949))).

63. *E.g.*, Wright, *supra* note 5, at 251-52; Feldman & Kim, *supra* note 43, at 531.

64. Hetcher, *supra* note 5, at 864; *see also* Keating, *supra* note 12, at 360-61 ("[T]he Hand Formula itself does not seem to dominate American negligence practice . . . .").

65. Hetcher, *supra* note 5, at 868.

66. Kelley, *supra* note 34, at 732 & n.4 (surveying tort casebooks and finding that "each casebook gives the *Carroll Towing Co.* formula a prominent place in its treatment of the standard of conduct in negligence cases").

67. Wright, *supra* note 5, at 252. Richard W. Wright and others have challenged this assumption. *Id.* at 253; Zipursky, *supra* note 42, at 2003 ("Breach in negligence law is to be judged by the ordinary care standard, and no evidence exists that either our system or the jurors who make the decisions are led to, or do, understand this standard in terms of the Hand Formula.").

68. *See, e.g.*, WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 23 (1987) ("People can apply the principles of economics intuitively—and thus 'do' economics without knowing they are doing it."); *see also supra* notes 49-51 and accompanying text (explaining how proponents of the Hand Formula argue that it guides decision-makers).

2. *Kantian Reasonableness*

While the Hand Formula is undoubtedly influential in the academy,[69] scholars have increasingly rejected the idea that reasonableness depends on, or should depend on, aggregate cost-benefit analysis.[70] Legal theorists have developed other accounts of reasonableness, with some of the most prominent grounded in Immanuel Kant's concept of equal freedom.[71]

Kantian accounts of reasonableness build upon Kant's assertion that "[f]reedom (independence from being constrained by another's choice), insofar as it can coexist with the freedom of every other in accordance with a universal law, is the only original right belonging to every man by virtue of his humanity."[72] The recognition of this innate freedom affords each individual human moral worth and dignity as a "self-legislating free rational being."[73] Preserving equal freedom requires one to honor others' autonomy rather than treat others as means to one's own ends.[74] From this perspective, "[a]ll persons should be treated as *ends in themselves* (i.e., as free and equal persons seeking to fully realize their humanity), rather than as mere means to be used to benefit others or society as a whole."[75]

Kantian legal theorists have argued that the concept of "equal freedom" informs tort law's reasonableness standard: to act reasonably is to act in a manner

---

69. Zipursky, *supra* note 42, at 1999 (referencing the "superstar status of the Hand Formula in torts").

70. *See* Eric Martínez & Kevin Tobia, *What Do Law Professors Believe About Law and the Legal Academy?*, 112 GEO. L.J. 111, 155 (2023) (finding that law professors responding to a survey largely rejected the notion that "efficiency" should inform reasonableness judgments).

71. *See, e.g.*, Wright, *supra* note 4, at 146-48, 211-23, 238; Weinrib, *supra* note 14, at 54. Theories in this vein have also sometimes been referred to as Aristotelian or Kantian-Aristotelian, and they are considered part of the family of corrective-justice theories. *See also* Andrew Fell, *Corrective Justice, Coherence, and Kantian Right*, 70 U. TORO. L.J. 40, 41 (2020) ("Therefore, Weinrib argues, corrective justice presupposes Kant's principle of right, in which the maintenance of equal freedom is fundamental."); GREGORY C. KEATING, REASONABLENESS AND RISK: RIGHT AND RESPONSIBILITY IN THE LAW OF TORTS 24 (2022) (describing corrective justice as "an ancient concept that has spawned a family of distinct modern conceptions"). Weinrib's Kantian conception of corrective justice, which emphasizes "'the unity of doing and suffering'—with the 'doing' being the infliction of the suffering by violating the 'abstract equality of free purposive beings under the Kantian conception of right,'" is focal in this Article, but "[o]ther theorists advance different conceptions of the wrong that corrective justice repairs." KEATING, *supra*, at 25 (quoting WEINRIB, THE IDEA OF PRIVATE LAW, *supra* note 17, at 58).

72. IMMANUEL KANT, THE METAPHYSICS OF MORALS 30 (Mary Gregor ed. & trans., Cambridge Univ. Press 1996) (1797) (emphasis omitted).

73. Wright, *supra* note 5, at 249.

74. KANT, *supra* note 72, at 30.

75. Wright, *supra* note 6, at 164-65.

that respects the equal freedom of others. For tort law, respecting equal freedom means avoiding inappropriately imposing risks on other people or their property.[76] But what makes risk imposition inappropriate? A workable Kantian perspective must find some way to balance "the conflicting liberties of injurers and victims: freedom of action (freedom to act and thereby impose risks) on the one hand, and security (freedom from accidental harm) on the other."[77]

Kantian tort theorists have struck this balance by distinguishing between "significant" and "insignificant" risks.[78] Imposing significant risks on others without their consent is inappropriate, while imposing insignificant risks on others is simply the background noise of day-to-day life. Framed in the language of reasonableness, it is unreasonable to impose significant risks on nonconsenting others; an actor who imposes such risks breaches his or her duty of care.[79] If the risk is realized and someone is injured, the actor will be liable to the injured party if the injured party can establish the other required elements of a negligence claim. Conversely, the imposition of insignificant risks is not a breach of the duty of care.[80] An actor whose conduct creates only insignificant risks will

---

76. *See, e.g.*, Arthur Ripstein, *Philosophy of Tort Law*, *in* OXFORD HANDBOOK OF JURISPRUDENCE & PHILOSOPHY OF LAW 656, 662-63 (Jules L. Coleman, Kenneth Einar Himma & Scott Shapiro eds., 2004) (noting that "not all risk imposition is inappropriate" and contending that "the boundary between appropriate and inappropriate risks is set by a system of norms of equal freedom").

77. Miller & Perry, *supra* note 1, at 351.

78. *See* Wright, *supra* note 5, at 256 ("[I]t is impermissible to use someone as a mere means to your ends by exposing him (or his resources) to significant foreseeable unaccepted risks, regardless of how greatly the benefit to you might outweigh the risk to him."); Ernest J. Weinrib, *Understanding Tort Law*, 23 VALPARAISO U. L. REV. 485, 518 (1989).

79. *See, e.g.*, Wright, *supra* note 5, at 256.

80. *See* Weinrib, *supra* note 78, at 518-19. Citing the British case *Bolton v. Stone* [1951] AC 850 (HL) 867 (appeal taken from Eng.), Weinrib contends that "[t]he common law allows action that does not expose potential sufferers to a *substantial risk*, understood as a function of the likelihood of the injury's occurring and the seriousness of the injury should it occur." *Id.* This *Bolton*-based approach—which uses *P* and *L* to distinguish "substantial risks," the imposition of which breaches the duty of care, from "insubstantial risks," the imposition of which does not breach the duty of care—reflects a strong Kantian view, leaving no room for *B* in the analysis. As a matter of British tort law, this approach was complicated by a subsequent case, *Wagon Mound No. 2*, which differentiated among three, rather than two, levels of risk a defendant's actions might create—far-fetched, fantastical risks; real but insignificant (or insubstantial) risks; and real and significant (or substantial) risks. Overseas Tankship (U.K.) Ltd. v. Miller Steamship Co. (*Wagon Mound No. 2*) [1967] AC 617 (PC) 641-42 (appeal taken from N.S.W.). This case opened the possibility that in the intermediate category of real but insignificant risks, a defendant may breach their duty of care if the risk was easily avoided (that is, if *B* was extraordinarily low). *Id.* This consideration of *B* in cases involving real but insubstantial risks is arguably inconsistent with Kantian theory; Weinrib certainly contended it was

not be liable for negligence.[81]

At first glance, categorizing risks as "significant" or "insignificant" seems to lack some of the ostensible certainty of the Hand Formula.[82] What makes a risk "significant"? Here, the Hand Formula inputs of *B*, *P*, and *L* help illuminate potential answers. Kantian theorists agree that any answer should involve some combination of the probability of the risk being realized (*P*) and the severity of expected harm if it is (*L*).[83] But Kantian theorists disagree with economic theorists as to whether and when the burden of mitigating the risk should play a role. That is, they disagree as to the role *B* plays in judging whether a party breached their duty of care. The unifying theme of Kantian theories is that the significance of the risks imposed by a party (based on some combination of the probability (*P*) and expected severity (*L*) of loss) is more important, relative to the burden (*B*) imposed by additional precautions, than the Hand Formula indicates.

Since breach is typically an issue for the jury, evaluating the state of the law of negligence requires a descriptive account of laypersons' definition of breach. In deciding whether litigants behaved reasonably, do laypeople like jurors weigh the burdens of taking precautions directly against the risks of proceeding without precautions, as economic theorists suggest? Or do they accord special weight to the significance of these risks, as Kantian theorists suggest?

The strongest Kantian views hold that the reasonableness of a litigant's conduct is entirely a function of the significance of the risks that conduct imposes, as reflected by some combination of the probability of harm and its expected gravity should it occur. These views leave no role for burden. Ernest Weinrib expressed one such strong view when he asserted that "the unreasonableness of the risk created by the defendant is seen in terms of the probability and gravity of its effects on others."[84] Where a risk is significant based on its probability and

a mistake. *See* WEINRIB, THE IDEA OF PRIVATE LAW, *supra* note 17, at 150 n.9. I argue that the *Bolton* analysis, using *P* and *L* to distinguish significant risks from insignificant risks, is a better reflection of Kantian reasonableness.

81. *See* Weinrib, *supra* note 78, at 518.

82. *See* Miller & Perry, *supra* note 1, at 355 (noting "a possible criticism of the Kantian [view] is that it fails to offer a clear criterion for distinguishing between reasonable and unreasonable conduct" – the economic view is "based on a comparison between two values," while the Kantian view focuses on "a single value that lies on an infinite continuum" without a well-defined comparator); *see also* Gregory Jay Hall, *Demystifying the Enigma: The Reasonable Person Standard in Tort*, 90 UMKC L. REV. 801, 827 (2022) ("Alternative approaches to negligence law only stand a chance of survival if they can provide a compelling theory of what the efficiency calculus centrally addresses, namely what, in particular, constitutes breach of the negligence standard of care.").

83. *See* WEINRIB, THE IDEA OF PRIVATE LAW, *supra* note 17, at 116.

84. *Id.*

gravity, "the cost of precaution is irrelevant."[85] Imposing a significant risk on a nonconsenting plaintiff is unreasonable (that is, a breach of the duty of care) because doing so deprives that plaintiff of autonomy, violating their equal freedom.

There are other views described as Kantian or Kant-influenced that do allow some limited role for burden (*B*), though they still afford *B* less weight relative to *P* and *L* than the Hand Formula does. One such example is the "disproportionate cost" perspective. On this view, conduct imposing risk on others is unreasonable if *B* is less than or modestly exceeds the product of *P* and *L*, but is reasonable if *B* far exceeds the product of *P* and *L*; the defendant need not incur a "disproportionate cost" to mitigate that risk.[86] Presenting another example, Wright has argued that proper Kantian reasoning takes burdens into account only in certain types of cases, and even then, they are assessed qualitatively rather than quantitatively.[87] For the purposes of this Article, I refer to these types of views as "softer" Kantian views. That label is not to say anything about their merits; rather, it is to express that they are more permissive with respect to the potential role of burden in reasonableness judgments.

Strong Kantian views leaving no place for burden yield clear, testable, a priori hypotheses about the criteria of reasonableness that cleanly contrast with those of economic views. Economic views predict that burdens matter; strong Kantian views predict they do not.[88] The contrast between economic and Kantian predictions becomes less clear if Kantian theory affords *B* a less determinate role. Therefore, I work from strong Kantian views to articulate the hypotheses for the experiments presented in Part II.[89] But while I frame my hypotheses in

---

85. *Id.* at 149.

86. *See* Gilles, *supra* note 17, at 513-14; *see also* Simons, *supra* note 8, at 1215-18 (describing an approach under which the analyst places greater weight on *PL* in "clear instances of unilateral risk imposition"). Indeed, the Kantian concern with significant, *unaccepted* risks may provide some room for *B* in such analyses, where risks may be deemed (implicitly) accepted if the benefits greatly outweigh the risks. While scholars have at times sought to reconcile the Kantian perspective with Hand-style balancing of costs and benefits, there has been debate about the extent to which such efforts at reconciliation preserve or undermine the Kantian idea of justice. For further analysis of this debate within the literature, see Wright, *supra* note 6, at 170-94.

87. *See* Wright, *supra* note 5, at 267.

88. As a result, strong economic views and strong Kantian views would lead to different outcomes in two types of cases: cases where the risk the defendant creates is significant but is nevertheless cost-justified, and cases where the risk the defendant creates is insignificant but is still not cost-justified. *See* Miller & Perry, *supra* note 1, at 353.

89. *See infra* Table 1 (laying out hypotheses grounded in strong Kantian views); *cf. infra* text accompanying note 270 (acknowledging that the Kantian effects observed in my studies might be reconcilable with softer Kantian views).

terms of strong Kantian views, my studies and their findings are relevant to any type of Kantian view. After all, any finding that decision-makers are influenced by the significance of risks (defined by some combination of *P* and *L*) *independent of* how those risks compare to the burden of precaution lends support to the fundamental Kantian notion that the significance of risks plays a larger role than the Hand Formula prescribes.[90]

The following Sections define two different strong Kantian views of reasonableness, which I refer to as broad and narrow. The left-hand side of Table 1 summarizes these views. Both are "strong" in that they define the significance of risks without regard to burdens, but they differ in the relative weight they accord to the probability of harm versus the expected gravity of harm should it occur. Specifically, the narrow view focuses primarily on the probability of harm (so long as its expected gravity is not de minimis).

90. *See infra* Section III.A.

TABLE 1. TAXONOMY OF KANTIAN VIEWS AS TO THE ROLES OF *B*, *P*, AND *L* IN JUDGING WHETHER CONDUCT IS REASONABLE

	Strong No Role for Burden (*B*)	**Softer** Some Role for Burden (*B*)
Broad *P* and *L* Get Equal Weight	Reasonableness depends on a combination of the probability of harm (*P*) and the expected severity of harm should harm occur (*L*). The burden of precaution (*B*) does not play a role in judging reasonableness.	Reasonableness depends primarily on a combination of the probability of harm (*P*) and the expected severity of harm should harm occur (*L*). The burden of precaution (*B*) may play a role in some cases, though it may be weighed less heavily than the Hand Formula prescribes.
Narrow *P* Gets More Weight than *L*	Reasonableness depends primarily on the probability of harm (*P*), where the expected severity of harm should harm occur (*L*) is not de minimis. The burden of precaution (*B*) does not play a role in judging reasonableness.	Reasonableness depends primarily on the probability of harm (*P*), where the expected severity of harm should harm occur (*L*) is not de minimis. The burden of precaution (*B*) may play a role in some cases, though it may be weighed less heavily than the Hand Formula prescribes.

a. Broad Kantian Reasonableness (Does PL *Exceed the "Significance" Threshold?)*

One plausible Kantian view is that the risks of harm are significant—and conduct creating those risks is unreasonable—when the product of the probability (*P*) and expected gravity (*L*) of harm exceeds some (non-*B*) threshold. In their influential article on the reasonable person, Alan D. Miller and Ronen Perry summarize Kantian perspectives as holding that "[t]he reasonableness of a risk is determined by its effect on human freedom, and the relevant criterion is the magnitude of the risk created by the potential injurer, namely the potential victim's expected harm."[91] More specifically, "[t]o determine the magnitude of the

91. Miller & Perry, *supra* note 1, at 351.

foreseeable risk created by the defendant, fact-finders need to consider the foreseeable severity of the harm and its foreseeable probability."[92] Miller and Perry's characterization captures the strong Kantian view asserted by Ernest Weinrib, who posited that whether conduct is reasonable is "seen in terms of the probability and the gravity of its effect on others."[93] In sum, these views focus on only one side of Hand's famous inequality: the probability (P) and expected severity (L) of harm influence whether conduct is judged reasonable, and this influence is entirely independent of the burden of precautions (B).[94]

Courts have at times embraced this broad Kantian view, most famously in the case of *Bolton v. Stone*.[95] The plaintiff in that case, Miss Stone, was standing on a road near her home when she was struck and injured by a stray cricket ball from a nearby cricket club.[96] The ball, launched from the club's field by an "exceptional" hit, carried a distance of one hundred yards, flying over a protective fence that was "some seventeen feet above the cricket pitch."[97] Miss Stone argued that the cricket-ground committee was negligent "in failing to take any sufficient precautions to prevent the escape of cricket balls from the ground and the consequent risk of injury to persons in [the road]."[98] Each of the law lords reviewing the case "assumed the defendant cricket club would be liable in negligence if the risk were foreseeable and of a sufficiently high level."[99] But each lord concluded that, on the particulars of the case, the risk "was not of a sufficiently high level to be deemed unreasonable as a matter of law, given the very low probabilities of a ball's being hit into the road and also hitting someone on the little-used residential side street."[100]

Though *Bolton v. Stone* did not explicitly refer to the Hand Formula, Lord Reid's opinion in *Bolton* neatly addressed his view of the relevance of each Hand Formula input. In considering whether the cricket club acted unreasonably, Lord Reid reasoned that

> it would be right to take into account not only how remote is the chance that a person might be struck [with a cricket ball] but also how serious

92. *Id.* at 353.

93. Weinrib, *Correlativity, Personality, and the Emerging Consensus on Corrective Justice*, *supra* note 17, at 116.

94. Weinrib, The Idea of Private Law, *supra* note 17, at 148.

95. [1951] AC 850 (HL) (appeal taken from Eng.).

96. *Id.* at 851.

97. *Id.*

98. *Id.* at 857.

99. Wright, *supra* note 5, at 261.

100. *Id.* at 261-62.

> the consequences are likely to be if a person is struck; but *I do not think it would be right to take into account the difficulty of remedial measures*.[101]

Indeed, from Lord Reid's perspective, if the cricket club could not operate without creating significant foreseeable risks for nonconsenting others, it should not operate at all—no matter how much benefit it provided for its owners and patrons.[102] Lord Reid's view of reasonableness was thus unconcerned with aggregate cost-benefit analysis: even if the risks of a stray cricket ball were entirely cost-justified, it would still be unreasonable for the cricket club to impose those risks on nonconsenting others like Miss Stone if the chance and seriousness of the risks were significant. In the language of the Hand Formula, Lord Reid viewed *P* and *L* as relevant, but *B* as irrelevant.

This Kantian perspective has also influenced American tort cases.[103] While courts have not often stated so plainly that they view the cost of precautions as irrelevant, scholars have at times posited that burden (*B*) plays a smaller role in real-world negligence cases than the significance of risks (*P* and *L*). This was the conclusion of Gary T. Schwartz, who, in a comprehensive study of New Hampshire and California tort opinions from the nineteenth century, found that courts "were reluctant to find that economic factors justified a defendant's risk-taking."[104] Indeed, he did not find a single case in which "mere monetary costs rendered nonnegligent a defendant's failure to adopt a particular safety precaution."[105] Similarly, John G. Fleming has observed that judicial opinions on negligence cases "do not make much of the cost factor."[106]

In sum, a broad Kantian perspective suggests that the combination of *P* and *L* dictates the significance of a particular risk, and thus whether an alleged tortfeasor acted unreasonably in imposing that risk on others. Critically, the evaluation of whether *P* and *L* create a significant risk is *not* made by weighing those factors against *B*. Instead, the combination of *P* and *L* is weighed against some other threshold. No scholarship from the Kantian perspective has yet attempted

---

101. *Bolton*, [1951] AC at 867 (emphasis added).

102. *See id.*

103. *See, e.g.*, Curtis v. State *ex rel.* Dep't of Transp., 180 Cal. Rptr. 843, 846, 855 (Ct. App. 1982) (affirming a jury verdict of negligence and reasoning that verdict was not inconsistent with the same jury's finding that, "as a whole," the benefits of defendant's conduct "outweighed the risks"); *see also* Wright, *supra* note 5, at 263 (arguing that "courts in the United States and elsewhere generally . . . find a defendant negligent if she created a foreseeable significant unaccepted risk of injury to the person or property of others," and collecting cases).

104. Gary T. Schwartz, *Tort Law and the Economy in Nineteenth-Century America*, 90 YALE L.J. 1717, 1757 (1980).

105. *Id.*

106. JOHN G. FLEMING, THE LAW OF TORTS 119 (8th ed. 1992).

to specify this threshold numerically, but the key point for present purposes is that whatever the threshold, it is independent of *B*.

b. Narrow Kantian Reasonableness (Where L *Is Not De Minimis, Does* P *Exceed the "Significance" Threshold?)*

From a Kantian perspective, it is unreasonable to impose "significant" risks on others without consent. As discussed above, one view is that whether a risk is "significant" is a function of both the probability of an accident (*P*) and its expected severity should it occur (*L*). But this is not the only plausible view of what distinguishes "significant" risks. Another plausible interpretation is that "significance" is primarily a function of *P*: risks are "significant" only when they are likely to materialize—that is, when *P* is high.

"Risk" has been defined as "the possibility of something bad happening."[107] This definition indicates that "risk" requires both a greater-than-zero *P* (a "possibility") and a greater-than-zero *L* ("of something bad"): there is no "possibility" of something when the probability of its occurrence (*P*) is zero, and any possibilities are not "of something bad" where the expected gravity of that something occurring (*L*) is zero. Working from this or a similar definition, attaching the modifier "significant" to "risk" can be understood to mean a *significant possibility* of something bad happening. This understanding is just as natural, if not more natural, than distributing the "significant" modifier across both the probability and severity aspects—that is, a *significant possibility* of something *significantly bad* happening. It also reflects the Kantian directive not to treat others as means to an end. For example, it seems evident that I treat others as means to an end when my actions, in pursuit of my own goals, impose a ninety-nine percent chance of a nonconsensual negative outcome arising for someone else. Whether that near-certain negative outcome is a bruise or death, I am nevertheless treating others as a means to my own ends.

The *P*-focused view of "significant risks" arguably finds support in a different law lord's opinion from *Bolton v. Stone*. Lord Oaksey noted that "[t]he standard of care in the law of negligence is the standard of an ordinarily careful man," before asserting that "an ordinarily careful man does not take precautions against every foreseeable risk."[108] Rather, "[h]e takes precautions against risks which are reasonably likely to happen."[109] The phrase "reasonably likely to happen" emphasizes the special import of *P*.

---

107. *Risk*, Cambridge Eng. Dictionary, https://dictionary.cambridge.org/us/dictionary/english/risk [https://perma.cc/5GEQ-LGN8].

108. Bolton v. Stone [1951] AC 850 (HL) 863 (appeal taken from Eng.).

109. *Id.*

Thus, both Kantian theory and case law provide plausible support for a narrow understanding of the criteria of reasonableness—one in which *P* is the weightiest Hand factor and drives outcomes so long as the severity of the prospective harm (*L*) exceeds some minimal threshold. Interestingly, a review of early torts scholarship reveals at least some theorizing that "negligence depend[s] alone on a 'probability' of injury."[110]

In sum, Kantians assert that the reasonableness inquiry turns on whether the defendant's conduct created significant risks for nonconsenting others, but the significance of those risks can plausibly be understood in broad (*PL*) or narrow (primarily *P*) terms.

C. The Empirical Study of Reasonableness

Section I.B described competing economic and Kantian theories of reasonableness. Section I.C.1 explains why it is important for both legal theorists and practitioners to understand which of these views best captures how laypeople judge reasonableness—and why experimental methods are ideal for this inquiry. Section I.C.2 reviews prior experimental work that has been done on lay reasonableness judgments, teeing up the unanswered Kantian question animating this Article's experiments: are laypeople judging tort disputes more influenced by information about the significance of risks (that is, by information about *P* and *L*) than the Hand Formula implies?

Before proceeding, it is worth emphasizing that the concept of reasonableness that is of interest here is the one used by lay jurors to categorize a defendant's behavior as "reasonable" or "unreasonable" in tort negligence cases. I refer to the concept of reasonableness used in this context as the "lay-legal concept" of reasonableness. This lay-legal concept may differ from laypeople's "ordinary," nonlegal concept of reasonableness,[111] and also from the "judicial" concept expressed by judges in their opinions.[112]

---

110. Green, *supra* note 33, at 1623, 1626 (citing Francis H. Bohlen, *The Probable or the Natural Consequences as the Test of Liability in Negligence*, 49 AM. L. REG. 79 (1901)) (discussing the early work of Francis H. Bohlen that focused on probability of harm).

111. *See* Tobia, *supra* note 6, at 86 (distinguishing "ordinary concepts" from "legal concepts"); Tobia, *supra* note 1, at 322, 324 (distinguishing between "reasonable" and "legally reasonable" based on whether study participants' reasonableness judgments were made in a legal frame); Julian Nyarko & Sarath Sanga, *A Statistical Test for Legal Interpretation: Theory and Applications*, 38 J.L. ECON. & ORG. 539, 540 (2021) (distinguishing between "Plain English" and "Legal English").

112. Nyarko & Sanga, *supra* note 111, at 540 (discussing Legal English and noting that "scholars have long suspected that judges apply their own understanding of *reasonable*, and that this understanding systematically diverges from the common [lay] understanding").

1. *Experimental Methods Can Provide Valuable Insights into the Lay-Legal Concept of Reasonableness*

The lay-legal concept of reasonableness merits careful attention for both practical and theoretical reasons. On a practical level, the issue of whether conduct is reasonable is typically decided by a jury, and with minimal instruction.[113] An improved understanding of how juries resolve this often-pivotal issue can help lawyers advise their clients, predict case outcomes, and craft more effective arguments.[114] On a theoretical level, proponents of both economic and Kantian views of reasonableness have claimed their views are descriptive, accounting for how decision-makers (typically lay jurors) evaluate reasonableness.[115] Experiments with lay participants are compelling methods for testing those empirical claims.[116]

Experiments afford researchers complete control over the information that participants see, allowing researchers to study whether changes in information they manipulate—called independent variables—result in changes in measured outcomes of interest—called dependent variables.[117] For instance, researchers can manipulate only certain aspects of negligence cases (such as information about the burdens of precautions (B) and the probability (P) and expected severity (L) of harms) to test, directly, how those manipulated aspects affect participants' reasonableness judgments.

To be sure, one could try to probe the relation between B, P, and L and lay-legal reasonableness judgments with other nonexperimental methods. For instance, one could review the often-mountainous records of real-world negligence cases and try to identify correlations between arguments touching on B, P, and L and jury verdicts. But countless confounding factors could potentially

---

113. *See, e.g.*, Gilles, *supra* note 10, at 814 ("[T]he operational meaning of negligence is largely determined by juries in particular cases . . . .").

114. Steven Hetcher, *The Jury's Out: Social Norms' Misunderstood Role in Negligence Law*, 91 GEO. L.J. 633, 633 (2003) (arguing that a literature on negligence that does not engage with lay reasonableness judgments unduly marginalizes the role of the jury and "badly skew[s] our understanding of the actual legal practices by which the negligence standard receives its content"). Indeed, ignoring the issue of how lay jurors judge negligence "run[s] afoul of the sound jurisprudential tenet that the best explanations of key legal concepts are those that provide a pragmatic explanation of the underlying legal practices." *Id.*

115. *See supra* notes 3-7 and accompanying text.

116. For an introduction to experimental techniques, see generally GEOFFREY KEPPEL & THOMAS D. WICKENS, DESIGN AND ANALYSIS: A RESEARCHER'S HANDBOOK (4th ed. 2004).

117. Russell Korobkin & Chris Guthrie, *Psychological Barriers to Litigation Settlement: An Experimental Approach*, 93 MICH. L. REV. 107, 119 (1994). Experimental research allows for causal inference because outcome differences across comparison groups cannot "be attributed to variables that were the same across comparison groups." *Id.* at 119 n.56.

influence jurors' decisions in any given case. By contrast, experimental techniques allow researchers to hold constant everything about their cases other than the manipulated variables, allowing researchers to cleanly infer that changes to the manipulated variables *caused* any corresponding changes in the outcome variable (that is, changes to information about *B*, *P*, and *L*, rather than any external factors, caused participants to change their reasonableness judgments).[118]

Other methods common in empirical legal research like surveying judicial opinions,[119] corpus linguistics,[120] and surveys of jury instructions[121] might be useful for investigating how judges or other legal experts understand reasonableness. Corpus linguistics (with the right corpus) can also shed light on the ordinary concept employed by laypeople in nonlegal contexts.[122] But these tools are not well suited to investigating the lay-legal concept.

Experimental methods also have notable advantages over other nonempirical approaches to legal philosophical questions like the nature of reasonableness. It is well established in the field of psychology that people often lack "direct introspective access to higher order cognitive processes."[123] When it comes to descriptive accounts of legal judgments, there may well be a limit on what even the most sophisticated theorists can learn through introspection.[124] By contrast,

---

118. *Id.*

119. *See, e.g.*, Schwartz, *supra* note 104, at 1718-19 (analyzing state-court tort-law decisions).

120. *See, e.g.*, Nyarko & Sanga, *supra* note 111, at 560 (using corpus-linguistics analyses and reporting significant differences between the judicial concept of reasonableness and the ordinary lay concept).

121. *See generally* Kelley & Wendt, *supra* note 27 (surveying jury instructions).

122. *See* Lucien Baumgartner & Markus Kneer, *The Meaning of 'Reasonable': Evidence from a Corpus-Linguistic Study*, *in* THE CAMBRIDGE HANDBOOK OF EXPERIMENTAL JURISPRUDENCE 440, 451-56 (Kevin Tobia ed., 2025) (providing corpus-linguistics data on lay nonlegal use of the word "reasonable").

123. Richard Nisbett & Timothy DeCamp Wilson, *Telling More Than We Can Know: Verbal Reports on Mental Processes*, 84 PSYCH. REV. 231, 231 (1977).

124. For further discussion of this issue, see Kevin Tobia, *Methodology and Innovation in Jurisprudence*, 123 COLUM. L. REV. 2483, 2492-95 (2023), which reviews JULIE DICKSON, ELUCIDATING LAW (2022). Indeed, there may be similar limits on what we can learn by directly asking other people to self-report their criteria for (un)reasonableness. *See* Neal R. Feigenson, *The Rhetoric of Torts: How Advocates Help Jurors Think About Causation, Reasonableness, and Responsibility*, 47 HASTINGS L.J. 61, 67-68 (1995). Psychologists often find that what people believe to be true about their behaviors and experiences is not empirically true. *See, e.g.*, Hayley J. Cullen, Helen M. Paterson, Timothy S. Dutton & Celine van Golde, *A Survey of What Legal Populations Believe and Know About Inattentional Blindness and Visual Detection*, 19 PLOS ONE, art. no. e0296489, at 1, 2 (2024) (discussing systematic "overestimation of visual detection, in that people believe that someone would be more likely to notice something than empirical research would suggest"); Ted Jaeger, *Circumscribed Shapes Are Enlarged: Is This a Variation of the Delboeuf Illusion?*, 104 PERCEPTUAL & MOTOR SKILLS 1116, 1116 (2007) (discussing optical illusions in which participants routinely overestimate the size of visual stimuli).

experimental methods examining judgments can uncover influences on those judgments that may not be obvious, or may even be counterintuitive, based solely on introspection.

2. *Experiments Examining the Lay-Legal Concept of Reasonableness*

Researchers conducting experiments to investigate lay-legal judgments of reasonableness ask participants to evaluate mock tort cases containing carefully manipulated information. By observing how the manipulated information affects participants' responses, researchers gain insight into the considerations driving lay reasonableness judgments.[125]

This Section focuses on the few experiments that have examined whether either economic or Kantian considerations influence lay reasonableness judgments. Only four prior papers have attempted to experimentally examine the role of economic considerations; they have done so by manipulating information related to (at least one of) the Hand Formula inputs of *B*, *P*, and *L* and documenting effects on lay judgments. No prior experiments have tested the Kantian proposition animating this Article—that the significance of the risks associated with conduct (reflected by the *PL* side of the Hand Formula) influences lay reasonableness judgments *independent* of the relation between those risks and the burden of taking additional precautions (*B*).

The earliest experiment to probe how Hand Formula inputs affect lay negligence judgments was reported by Edward Green in 1967.[126] Green asked 192 participants to render individual judgments of "liable" or "not liable" in a negligence case brought on behalf of a child who fell into the defendant homeowner's

---

125. *See generally* Christopher Brett Jaeger, *Reasonableness from an Experimental Jurisprudence Perspective*, *in* THE CAMBRIDGE HANDBOOK OF EXPERIMENTAL JURISPRUDENCE, *supra* note 122, at 428-39 (surveying prominent theories about the considerations influencing reasonableness). Some of these considerations may be reflected in legal theory (such as *B*, *P*, and *L*). Other considerations may be wholly inconsistent with legal theory. For example, experimental findings indicate that laypeople evaluating negligence are influenced by mental-state information about the defendant, though tort theory indicates such information should be irrelevant. Francesco Margoni & Teneille R. Brown, *Jurors Use Mental State Information to Assess Breach in Negligence Cases*, 236 COGNITION, art. no. 105422, at 1, 7 (2023). In addition, empirical findings indicate that lay judgments about the reasonableness of precautions are outcome-dependent, even though legal theory holds such judgments should be outcome-independent. *See* Markus Kneer, *Reasonableness on the Clapham Omnibus: Exploring the Outcome-Sensitive Folk Concept of Reasonable*, *in* JUDICIAL DECISION-MAKING: INTEGRATING EMPIRICAL AND THEORETICAL PERSPECTIVES 25, 36 (Piotr Bystranowski, Bartosz Janik & Maciej Próchnicki eds., 2022); Kim A. Kamin & Jeffrey J. Rachlinski, *Ex Post ≠ Ex Ante: Determining Liability in Hindsight*, 19 LAW & HUM. BEHAV. 89, 99 (1995).

126. Edward Green, *The Reasonable Man: Legal Fiction or Psychosocial Reality?*, 2 LAW & SOC'Y REV. 241, 242 (1967).

unattended backyard swimming pool.[127] Green experimentally manipulated four aspects of the case: (1) whether the fence the homeowner put around the pool was three or six feet high; (2) whether the home was located in a neighborhood with a "noticeably large" or "noticeably small" proportion of preschool and elementary-school children; (3) whether the child who fell into the pool experienced minor injury or death; and (4) whether the "jury instructions" given to participants were short and general, or longer and more explicitly framed in terms of the Hand Formula.[128] Fence height produced a noticeable effect on lay judgments: participants were significantly less likely to find the defendant liable for negligence when the fence was six, rather than three, feet tall.[129] More relevant for present purposes, Green also found that the proportion of children in the neighborhood (relating to *P*, in Hand Formula terms) significantly influenced verdicts among the subset of participants in the three-foot-fence condition.[130] (Green did not manipulate *B* or *L*.[131])

In 1988, Marylie Karlovac and John M. Darley investigated the negligence judgments of 236 undergraduate participants using similar methods.[132] Karlovac and Darley presented participants with a case involving a man who precariously parked a heavy construction buggy in a sloped alleyway, manipulating two things about the case: (1) the number of recommended precautions the man took to keep the buggy from rolling downhill and (2) the severity of expected losses were the buggy to roll away.[133] In most experiments, participants were told the buggy rolled down the hill and hit either a low-severity or moderate-severity target.[134] Participants rated, on a nine-point Likert scale, the extent to which the man's behavior with the buggy was negligent. Echoing Green's findings, Karlovac and Darley found that increasing the level of precaution taken led

127. *Id.* at 243-44. The sample included a combination of students and nearby community members. *Id.* at 244.

128. *Id.* at 244.

129. *Id.* at 245.

130. *Id.* at 246.

131. *See infra* note 138. Green manipulated the injury *actually suffered* by the child, not the (ex ante) expected severity of injury should an injury occur, which is the *L* of interest under the Hand Formula and on Kantian views. *See* JENNIFER K. ROBBENNOLT & VALERIE P. HANS, THE PSYCHOLOGY OF TORT LAW 56 (2016) ("[I]t is the harm that is risked, and not the harm that results, that is relevant to the analysis of whether the actor acted with reasonable care."). For a review of empirical work on how the injury a plaintiff actually suffers affects legal judgments, see generally Jennifer K. Robbennolt, *Outcome Severity and Judgments of "Responsibility": A Meta-Analytic Review*, 30 J. APPLIED SOC. PSYCH. 2575 (2000).

132. Marylie Karlovac & John M. Darley, *Attribution of Responsibility for Accidents: A Negligence Law Analogy*, 6 SOC. COGNITION 287, 287 (1988).

133. *Id.* at 291-92, 298.

134. *Id.* at 292, 298.

participants to rate the defendant as less negligent.[135] Karlovac and Darley also generally found that reducing the severity of expected loss in the event of an accident (*L*, in Hand Formula terms) reduced lay negligence ratings, at least at some levels of precaution.[136] (Karlovac and Darley did not manipulate *B* or *P*.[137])

The experiments by Green and by Karlovac and Darley were inspired by the Hand Formula, but they did not—and could not—test whether participants' ultimate judgments were consistent with the Hand Formula. None of the experiments assigned concrete values for *B*, *P*, and *L*.[138] Thus, in each experiment, it was ambiguous whether the defendant failed to take a cost-justified precaution[139]—and therefore impossible to tell whether observed influences of probability of harm (*P*) and gravity (*L*) reflected economic or Kantian reasoning.

My own 2021 paper included two experiments in which participants received unambiguous Hand Formula information.[140] I presented 210 participants with written vignettes describing a variety of negligence cases.[141] For each case, participants received full *BPL* information related to a specific untaken precaution. I manipulated whether the untaken precautions were cost-justified ($B < PL$) or were not cost-justified ($B > PL$), but I only tested these two possibilities—I did not independently manipulate *P* or *L* information to see if participants gave it extra Kantian weight.[142] Further, participants in those experiments also received information about whether the untaken precaution was *customary*, in addition to *BPL* information. I found that participants' negligence judgments were affected by custom information (information about the percentage of people who would have taken the precaution the defendant omitted) and not *BPL* information.[143] But, again, these studies tested whether participants would use *BPL* information

135. *Id.* at 295, 303, 306, 309-10, 313.

136. *Id.* It was at low or moderate levels of precaution that negligence judgments tended to reflect a concern for the harms being risked. *Id.* at 295, 305-06, 310-11.

137. *Id.* at 296.

138. Green's experiment did not explicitly identify any specific untaken precaution to which the Hand Formula could apply; participants were not told that the plaintiff was arguing that a taller fence would have reduced the risk or prevented the injury (nor were they told the costs of building a taller fence). Green, *supra* note 126, at 244. Karlovac and Darley's studies did explicitly identify untaken precautions in some conditions, but the burden associated with these precautions (*B*) was never expressed on a common scale with the expected losses those precautions would prevent (*PL*). Karlovac & Darley, *supra* note 132, at 291-92.

139. *See* Gilles, *supra* note 4, at 1025 (arguing that the Hand Formula treats negligence as tantamount to "omitt[ing] a cost-justified precaution").

140. Christopher Brett Jaeger, *The Empirical Reasonable Person*, 72 ALA. L. REV. 887, 911-22 (2021).

141. *Id.*

142. *Id.* at 911, 913-14, 919.

143. *Id.* at 916-22.

in the economic manner that the Hand Formula predicts. Further, it is quite possible that the influence of custom information overpowered any role that economic considerations would have otherwise played, necessitating future studies isolating *BPL* for analysis.[144]

Most recently, Shoham Choshen-Hillel, Ehud Guttel, and Alon Harel conducted a series of similar mock-juror experiments.[145] Participants judged whether defendants who engaged in a risky action, either to obtain a gain or prevent a loss, acted reasonably. For example, in one scenario, a defendant farmer sprayed their trees either to increase their yield by $5,000 worth of fruit or eliminate pests that would otherwise cause a loss of $5,000 worth of fruit.[146] But the spray could harm trees on neighboring farms, and the benefits (or loss savings) of the defendant farmer spraying ($5,000) could be greater or less than the expected losses imposed on the neighboring farmers ($100 or $7,000).[147] Participants were more likely to find a defendant's behavior unreasonable—and that the defendant was therefore liable for negligence—when the defendant's behavior was intended to obtain a gain, as opposed to prevent a loss.[148] More importantly for present purposes, however, participants were also more likely to find the defendant's choice to spray unreasonable when it was not cost-justified.[149] In other words, Choshen-Hillel, Guttel, and Harel found evidence that lay negligence judgments *are* influenced by cost justification under the Hand Formula—at least when the participants are presented with Hand Formula information and *not* custom information.[150] But Choshen-Hillel and colleagues'

144. Another possibility is that the wording of the (representative) jury instructions used in these experiments affected the criteria participants used in their judgments. *See* Kevin Tobia et al., *The Nature of Reasonableness*, 61 STAN. J. INT'L L. 111, 128 (2025) (raising this possibility). However, empirical research on jury instructions suggests that the precise wording was unlikely to meaningfully influence participants. *See* Jaeger, *supra* note 140, at 938 n.277 (citing research on the limited influence of jury instructions).

145. Shoham Choshen-Hillel, Ehud Guttel & Alon Harel, *Framing Negligence*, 19 J. EMPIRICAL LEGAL STUD. 296, 300-02 (2022).

146. *Id.* at 333-34.

147. *Id.*

148. *Id.* at 297, 320.

149. *Id.* at 304; *see also id.* at 308-09, 311, 315 (demonstrating that this tendency was observed across multiple scenarios). The authors also found evidence that the cost efficiency of defendants' behavior was more influential in loss conditions than in gain conditions. *Id.* at 320.

150. *See id.* at 320; *cf.* W. Kip Viscusi, *Punitive Damages: How Jurors Fail to Promote Efficiency*, 39 HARV. J. LEGIS. 139, 144, 158-59 (2002) (observing, based on case studies and experimental findings, that risk analysis can become a negative feature in jury determinations of punitive damages for corporate defendants, even where the defendant proceeded efficiently); W. Kip Viscusi, *Corporate Risk Analysis: A Reckless Act?*, 52 STAN. L. REV. 547, 550-52 (1999)

experimental manipulations of *BPL* information were binary—precautions were cost-justified under the Hand Formula or they were not—so their findings do not address the relative weight that lay participants put on the different components of the Hand Formula or the Kantian hypotheses animating this Article.

In sum, surprisingly little experimental work has directly tested the economic hypothesis that actions are judged unreasonable when risks outweigh burdens and reasonable when burdens outweigh risks. Further, no experimental work to date has tested the Kantian hypothesis that reasonableness judgments are influenced by the significance of risks independent of their relation to burdens. Part II reports the first experiments in this vein.

II. THE EXPERIMENTS: WHICH HAND FORMULA INPUTS DRIVE LAY REASONABLENESS JUDGMENTS?

This Part reports five original experiments testing how information about the Hand Formula inputs affects lay reasonableness judgments. The experiments rely on similar methods and logic. In each, participants acted as mock jurors, evaluating a tort negligence case for which they were given complete information about the burden of taking an additional precaution (*B*), the increased probability of harm occurring in the absence of that precaution (*P*), and the expected severity of the harm should it arise (*L*). *B*, *P*, and *L* values were manipulated to test the economic hypothesis that lay negligence judgments are influenced by the relation between *B* and *PL* in accordance with the Hand Formula, and the Kantian hypothesis that the risk (*PL*) side of the Hand Formula has independent influence on lay negligence judgments beyond what the Hand Formula predicts.

My results provide support for both hypotheses. Participants found defendants' conduct more unreasonable when *B* was less than *PL*, as one would expect if participants engaged in some degree of Hand-style cost-benefit analysis. But, critically, information about the significance of risks (about *P* and *L*) affected participants' reasonableness judgments more than the Hand Formula suggests: participants found defendants' conduct more unreasonable as the value of *PL*—and particularly the value of *P*—increased, whether *B* was less than *PL* or not. In short, the Hand Formula captures something about how laypeople judge reasonableness, but it fails to capture the Kantian weight laypeople put on the significance of risks.

(presenting similar findings). While Viscusi's studies did not directly test negligence judgments, they revealed a willingness of lay study participants to impose punitive damages for cost-justified conduct.

Section II.A describes the general methodology and experimental logic common to all five experiments. Sections II.B through II.F then provide the details of each experiment and its findings.

A. General Methods and Logic

1. General Methods

Participants[151] in each of the five experiments experienced a similar sequence of events. Participants read a vignette describing a tort negligence case and judged whether the defendant acted unreasonably.[152] In each case, a defendant imposed a risk on a nonconsenting plaintiff while pursuing the defendant's own private goals.[153] When the risk materialized, the injured plaintiff claimed the

---

151. Participants in all five experiments were Americans of at least eighteen years of age. They were recruited through Prolific, an online platform frequently used by researchers. *See, e.g.*, Krin Irvine, David A. Hoffman & Tess Wilkinson-Ryan, *Law and Psychology Grows Up, Goes Online, and Replicates*, 15 J. EMPIRICAL LEGAL STUD. 320, 343 (2018) (noting that Prolific "was built to serve academic audiences"); Benjamin D. Douglas, Patrick J. Ewell & Markus Brauer, *Data Quality in Online Human-Subjects Research: Comparisons Between MTurk, Prolific, CloudResearch, Qualtrics, and SONA*, 18 PLOS ONE, art. no. e0279720, at 4, 12 (2023) (identifying Prolific as one of the "most commonly used online data collection platforms" and reporting evidence that Prolific provides high-quality data relative to other platforms). Prolific directed participants to a Qualtrics survey that contained the experiment. Participants in all studies provided informed consent and were paid $1.40 for completing Experiments One, Four, and Five and $1.60 for completing Experiments Two and Three (which had slightly longer vignettes). These amounts were calculated to pay a minimum rate of $12 per hour. Each participant could only complete one of the five experiments. Demographic information about the participants in each of the five studies is included in the Appendix.

All experiments were approved as exempt by the Baylor University Institutional Review Board (IRB) and preregistered through the Open Science Framework (OSF). IRB letters approving each study as exempt are uploaded to OSF. For the general OSF for this project, see Christopher Brett Jaeger, *The Hand Formula's Unequal Inputs*, OSF (July 24, 2025), https://osf.io/gfnu4/?view_only=bb29d8f003854756a4f1b84c3ed7d5a7 [https://perma.cc/KV3F-WBHF]. I preregistered for each experiment, *inter alia*, my hypotheses, sample sizes (based on a priori power analyses), planned analyses, and criteria for excluding participants. Study materials and anonymized data files are available through OSF.

152. Each experiment employed a between-subjects design, meaning every participant evaluated only one case under one experimental condition (that is, one combination of *B*, *P*, and *L* information).

153. These contours of the cases are noteworthy for reasons related to one of the softer Kantian views discussed in Section I.B.2, *supra*. Kantian theorist Richard W. Wright allows that *B* may properly factor into reasonableness judgments in some types of cases—but *not* those where the defendant puts the plaintiff at risk for the benefit of the defendant or some third party (with no direct or indirect benefit to the plaintiff). *See* Wright, *supra* note 5, at 261-63. My experiments were designed to involve such cases, where strong Kantian views and Wright's more nuanced view agree that *P* and *L* should matter, while *B* should not.

defendant was negligent for failing to take a particular precaution that would have prevented the plaintiff's injury. Each case included manipulated information about the cost of taking the untaken precaution (*B*), the increase in the probability of harm arising without the precaution versus with the precaution (*P*), and the expected severity of the harm should it occur (*L*),[154] all in accessible language.[155]

The case used in Experiment One, for example, involved a dispute between two neighboring farmers. The plaintiff claimed that their uphill neighbor's negligent choice of farming methods caused a landslide that damaged their property. The plaintiff asserted that the defendant should have used a safer (but more expensive) farming method that would have eliminated the landslide risk. In this vignette, a short *B* paragraph stated how much more the defendant farmer would have had to spend to use the safer method, a *P* paragraph stated the probability of a landslide with the riskier farming method the defendant used (and that this was relative to zero risk with the safer method), and an *L* paragraph stated the value of the property threatened by the landslide.[156] The order in which the *B*, *P*, and *L* paragraphs appeared was counterbalanced.[157]

At the end of the case, participants read that the plaintiff sued the defendant, claiming that the defendant's failure to take the specified precaution was negligent. Brief, neutral jury instructions tasked participants with deciding whether the defendant was negligent, specifying that the defendant was negligent if they acted unreasonably by failing to take the relevant precaution. The instructions

---

154. Specifically, the manipulated *B*, *P*, and *L* information was relayed to participants through three short paragraphs in the middle of the case, with one short paragraph corresponding to each input.

155. Case vignettes were written at seventh- to ninth-grade reading levels, with Flesch Readability Ease scores in the sixties and seventies. *See Flesch Kincaid Calculator*, GOOD CALCULATORS, https://goodcalculators.com/flesch-kincaid-calculator [https://perma.cc/E7C6-D9HD] (measuring the comprehensibility of documents, with a Flesch Readability Ease score around seventy to eighty equal to a seventh-grade reading level). Participants reported that they did not find study materials confusing. When asked whether they found the cases and prompts confusing, participants' mean responses never exceeded 1.53 on the scale from 1 (Not At All Confusing) to 9 (Very Confusing).

156. To help ensure participants did not overlook the manipulated information, the *B*, *P*, and *L* paragraphs were always immediately followed by a bolded summary paragraph. This paragraph stated that "[t]o summarize, the following facts are true," and then presented three bullet points reiterating the *B*, *P*, and *L* values (in the same order they were presented in the *B*, *P*, and *L* paragraphs).

157. Specifically, the order was fully counterbalanced in the first four experiments. Statistical tests indicated that the order in which material was presented had no systematic effects on participants' negligence ratings in Experiments One through Four. Given that there were no order effects in the first four experiments, the order was not counterbalanced in Experiment Five.

also clarified that the court would order the defendant to pay for the plaintiff's losses only if the defendant was negligent.

Beneath these jury instructions, participants were asked to rate, on a Likert scale from one (Completely Disagree) to eight (Completely Agree), the statement that "[Defendant] was negligent, meaning [Defendant] acted unreasonably by [proceeding without the relevant precaution]."[158] These negligence ratings served as my primary dependent variable in all experiments.[159]

After providing a negligence rating and any explanation for their decisions, participants were to rate the extent to which information regarding *B*, *P*, and *L* individually played a role in their decision on a nine-point Likert scale from one (No Role At All) to nine (Major Role).[160] These ratings allowed me to probe the secondary question of whether participants' stated reports of what influenced their judgments reflected what in fact influenced their judgments; psychological evidence suggests that people are often not conscious of influences on their judgments.[161]

2. *Experimental Logic*

The details of the case vignettes and the values of *B*, *P*, and *L* used in each experiment are described in Sections II.B to II.F. Across all five experiments, two important principles guided the selection of *B*, *P*, and *L* values.

First, I assigned the *B*, *P*, and *L* values in every experiment such that the low value of *B* was lower than the lowest possible product of *P* and *L*, and the high value of *B* was higher than the highest possible product of *P* and *L*. Ensuring that the low *B* value was always lower than *PL* and the high *B* value was always

158. Likert scales are commonly used in empirical legal research, and social-science research generally, because they are more sensitive than binary responses. The Likert scale I use parallels that used by Francesco Margoni and Teneille R. Brown, but with no "neutral" midpoint. *See* Margoni & Brown, *supra* note 125, at 5. The lack of any neutral option allowed me to recode participants' responses as binary verdicts of "not negligent" (negligence ratings one through four) or "negligent" (negligence ratings five through eight) and evaluate whether my manipulations affected those binary verdicts. I found that the effects of *B*, *P*, and *L* information on binary verdicts largely paralleled the effects of *B*, *P*, and *L* information on eight-point negligence ratings. *See infra* note 266 and accompanying text.

159. *See* Korobkin & Guthrie, *supra* note 117, at 119 (discussing potential issues with using a binary win/loss metric in analyzing case outcomes). When providing negligence ratings, participants also had the opportunity to explain why they decided as they did.

160. Each prompt reprinted the relevant *B*, *P*, or *L* information verbatim from the bolded summary paragraph in the case to prevent any confusion as to what participants were rating. These prompts appeared in the same order the information appeared in the case. Participants could not return to prior screens.

161. Nisbett & Wilson, *supra* note 123, at 231-33.

higher than *PL* made the economic and Kantian hypotheses distinct and testable. Effects of *B* would indicate that participants were influenced by economic considerations, and effects of *P* and *L* would indicate that participants were influenced by Kantian considerations.[162]

The second guiding principle for assigning *B*, *P*, and *L* values was that, in each experiment, I sought to include among the possible combinations of *B*, *P*, and *L* the two situations for which economic and Kantian views yield clearly opposing predictions: cases where *PL* is too low to be "significant" but *B* is even lower, and cases where *PL* is high enough to be "significant" but *B* is even higher.[163] To present these situations, I ensured the low *B*s in my experiment were quite low (from $4 to $90) and the high *B*s were quite high (from $9,000 to $100,000). The use of rather extreme values in my *B* manipulation also strengthens the test of the Kantian hypotheses. These *B* values were designed to capture attention and maximize any influence *B* may have on reasonableness judgments. If participants' judgments are nevertheless influenced by *P* and *L* independent of *B*, it strongly suggests their judgments are influenced (at least in part) by Kantian considerations.

I must comment briefly on the fact that all five experiments involve cases for which participants have complete *B*, *P*, and *L* information. It is, of course, rare that jurors have complete, unambiguous *B*, *P*, and *L* information in real litigation.[164] But this does not limit my studies. Neither economic nor Kantian theories of reasonableness limit themselves to cases with perfect information; both aim to explain something deeper and more general about the nature of

---

162. Specifically, assigning values in this way made the *B* variable completely synonymous with cost justification under the Hand Formula: when *B* was low, the precaution the defendant failed to take was always cost-justified (i.e., $B < PL$), and when *B* was high, it was never cost-justified (i.e., $B > PL$). If participants judge reasonableness based solely on the Hand Formula, I should therefore observe effects of *B*—participants should rate defendants more negligent when *B* is low than when *B* is high—and no effects of *P* or *L*. In contrast, Kantian views would predict that the reasonableness of defendants' conduct depends on some combination of *P* and *L* independent of their relation to *B*. Given my manipulations, any significant effects of *P* or *L* would provide evidence of such an independent Kantian influence. Recall that broad Kantian views consider both *P* and *L*, whereas narrow Kantian views focus on *P* as long as *L* is not de minimis. If participants judge reasonableness based on broad Kantian considerations, I should observe higher negligence ratings when *P* is high than when *P* is low and when *L* is high than when *L* is low. If participants judge reasonableness on narrow Kantian grounds, I should observe effects of *P*—participants should rate the defendant more negligent when *P* is high than when *P* is low—and no effect of *L*. If participants judge reasonableness purely based on what I have called strong Kantian views, whether narrow or broad, there would be no effect of *B* at all.

163. Miller & Perry, *supra* note 1, at 353-54.

164. *But see* R.I. Hosp. Tr. Nat'l Bank v. Zapata Corp., 848 F.2d 291, 294-95 (1st Cir. 1988) (providing an example of when evidence allowed a court to quantify and compare *B* against *PL*).

negligence. If (as these theories suggest) people judging reasonableness truly seek to evaluate cost justification or the significance of risks in the absence of full information, they should certainly do so in the presence of full information.[165]

Thus, "observing how [*B*, *P*, and *L*] information affects laypeople when it is available and unambiguous provides insight as to what laypeople consider relevant when evaluating negligence."[166] If participants' negligence judgments vary depending on whether *B* is greater or less than *PL*, it would suggest that the Hand Formula captures some truth about the lay-legal concept of reasonableness.[167] If participants' judgments are influenced by *P* or the combination of *P* and *L*, independently of those factors' roles in the Hand Formula, it would suggest that Kantian perspectives capture some truth about the lay-legal concept of reasonableness.[168]

Experiments One, Two, and Three provided basic tests of economic and Kantian hypotheses in different types of cases: a property-damage case (Experiment One), a personal-injury case involving an individual defendant (Experiment Two), and a personal-injury case involving a corporate defendant (Experiment Three). Participants in these initial experiments were influenced to some degree by the balance of costs and benefits under the Hand Formula, at least where the defendant was an individual rather than a corporation.[169] But participants were also influenced by an independent Kantian concern with the significance of risk. In all three experiments, participants' negligence ratings increased as the value of *PL*—and, particularly, the value of *P*—increased, without regard to the value of *B*.

Experiment Four revisited the property-damage case used in Experiment One with new *B*, *P*, and *L* values, finding once again that both economic and Kantian concerns influenced participants' judgments. Experiment Five then replicated Experiment Four, with all the relevant *BPL* math carefully laid out for participants; the case vignette specifically told participants that cost-benefit analysis indicated the defendant should (or should not) have taken the relevant

---

165. *See* McCarty v. Pheasant Run, Inc., 826 F.2d 1554, 1557 (7th Cir. 1987) (lamenting that juries are "forced to make rough judgments of reasonableness, intuiting rather than measuring the factors in the Hand Formula," because parties "do not give the jury the information required to quantify" those factors).

166. Jaeger, *supra* note 140, at 910.

167. *See supra* Section I.B.1.

168. *See supra* Section I.B.2. To be clear, I did not expect that either an economic or Kantian view, standing alone, would capture everything about lay reasonableness judgments. *See* GEORGE P. FLETCHER, TORT LIABILITY FOR HUMAN RIGHTS ABUSES 103 (2008) ("There is . . . much vagueness in trying to compute the factors that enter into assessing whether risks are reasonable or not.").

169. For further discussion of this point, see *infra* notes 280-281 and accompanying text.

precaution. Thus, Experiment Five ruled out the possibility that the Kantian effects observed in the first four experiments were due to math aversion or errors.

B. Experiment One: The Influence of B, P*, and* L *in a Property-Damage Case*

My first experiment tested economic and Kantian hypotheses about the roles of *B*, *P*, and *L* in a property-damage case.[170] The results supported both hypotheses. Participants found the defendant's actions to be more unreasonable where $B < PL$, consistent with some degree of Hand-style comparison of costs and benefits. But participants' judgments were also independently influenced by *PL*—particularly *P*—indicating that Kantian considerations had an effect.

1. The Case

Dylan, a private farmer, faces a choice between two farming methods: the safer, more expensive Method A and the riskier, less expensive Method B. Either method would produce the same quality and quantity of harvest for Dylan. But the riskier, less expensive Method B entailed—even if perfectly executed—some risk of a landslide that would destroy the barn of Dylan's neighbor Pat, which was located downhill from Dylan's property. Method A entailed no such risk. Dylan, aware of the relevant costs and benefits, opted for the riskier, less expensive Method B. Though Dylan executed Method B perfectly, a landslide from Dylan's property destroyed Pat's barn, and Pat sued Dylan for negligence.

Participants read about this case in a written vignette of about 350 words.[171] The vignette included values for all three Hand Formula inputs as they related to Dylan's choice of farming method. The *B* paragraph specified the additional cost of the safer, more expensive farming method. The *P* paragraph stated the probability of a landslide with the riskier, less expensive farming method and stated that there was zero risk of a landslide with the safer, more expensive method. The *L* paragraph specified (1) that Pat's barn would certainly be

---

170. For an introduction to experimental techniques, see generally KEPPEL & WICKENS, *supra* note 116. Researchers use experimental designs to study whether changes in variables they manipulate—called independent variables—result in changes in measured outcomes of interest—called dependent variables. The chief advantage of experimental designs is that experimenters can "hold constant the variables they are not testing." Korobkin & Guthrie, *supra* note 117, at 119. Experimental research allows for causal inference because outcome differences across comparison groups cannot "be attributed to variables that were the same across comparison groups." *Id.* at 119 n.56.

171. The full vignette is available on OSF. *See* Christopher Brett Jaeger, *Hand Formula's Unequal Inputs – Vignettes Used in Experiments*, OSF (July 24, 2025), https://osf.io/ta85j/files/feacd?view_only=91e8899ee35f45dd8423234973647162 [https://perma.cc/2M4B-CESL].

destroyed in the event of a landslide; (2) that nothing else of any value would be at risk in a landslide; and (3) the features and precise value of the barn.

I manipulated the *B*, *P*, and *L* values independently in a 2x2x2 factorial design.[172] For *B*, the safe Method A could be either $6 or $100,000 more expensive than the riskier Method B. For *P*, the risk of a landslide could be either 0.2% or 20% with the riskier Method B, as opposed to zero with Method A. For *L*, the value of Pat's barn—certain to be destroyed should a landslide occur—could be either $4,000 or $400,000.[173] With these values assigned to *B*, *P*, and *L*, any effect of *B* on participants' evaluation of the case would indicate that participants were influenced by economic considerations. Any effects of *P* and *L* would indicate that participants were influenced by Kantian considerations, weighing the significance of risks more heavily than economic perspectives suggest.

Recall that participants evaluated the case by providing a negligence rating. Here, this meant rating their agreement with the statement that "Dylan was negligent if Dylan acted unreasonably in choosing to farm with Method B rather than Method A" on an eight-point Likert scale. Subsequently, participants rated the extent to which they believed information about *B*, *P*, and *L* played a role in their decision.[174]

2. *Results*

Participants' negligence ratings[175] were significantly influenced by *B* (the additional cost of the safer farming method) and *P* (the increased probability of a landslide with the riskier farming method), but not by *L* (the expected injury in the event of a landslide).[176] These findings suggest that both economic and

---

172. *See infra* Appendix Table 1.

173. Notably, the high *B* value exceeds the highest possible product *PL* by the same proportion that the lowest possible product *PL* exceeds the low *B* value (rounding to the nearest dollar). In addition, with the values chosen, the proportion of the high *P* value to the low *P* value is the same as the proportion of the high *L* value to the low *L* value.

174. Participants could not return to prior screens.

175. N = 214. Seven participants were excluded from analysis based on a priori, preregistered exclusion criteria. Note that I opened all five experiments on Prolific seeking 220 participants. Due to interactions between the Prolific platform that I used to recruit participants and the Qualtrics survey used to administer my experiments, I sometimes ended up with 219 or 221 total responses before applying my exclusion criteria.

176. My primary preregistered analysis was a three-way factorial analysis of variance (ANOVA) assessing the influence of my manipulations of *B*, *P*, and *L* on participants' negligence ratings. "Analysis of variance (ANOVA) is a statistical technique used to evaluate the size of the difference between sets of scores." GLENN GAMST, LAWRENCE S. MEYERS & A.J. GUARINO, ANALYSIS OF VARIANCE DESIGNS: A CONCEPTUAL AND COMPUTATIONAL APPROACH

Kantian views (particularly a narrow Kantian view) capture something about lay reasonableness judgments. On the economic side, the effect of *B* is consistent with participants engaging in some balancing of costs against benefits. But, on the (narrow) Kantian side, the effect of *P* indicates that *P* information influenced participants above and beyond its role in cost-benefit balancing. This effect is incompatible with pure Hand Formula reasoning.[177]

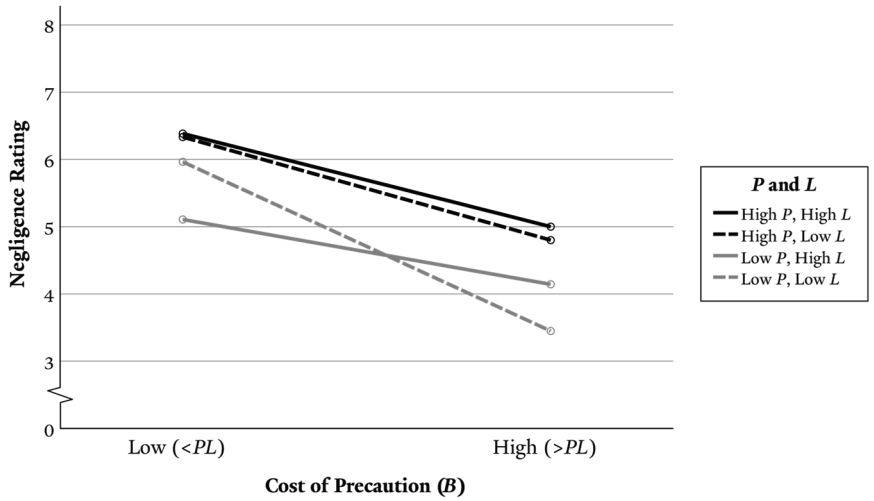
Figure 1 plots the means of participants' negligence ratings for each possible combination of *B*, *P*, and *L*.[178]

WITH SPSS AND SAS 3 (2008). ANOVA is among the most common analyses in psychological research. There has been some history of interdisciplinary debate about the use of ANOVA versus nonparametric tests for analyzing Likert-scale data, but that debate is immaterial to this Article: for all experiments, use of nonparametric Kruskal-Wallis tests yields the same patterns of results obtained with my primary ANOVA analyses. Statistical analyses in my experiments were conducted with IBM SPSS Statistics Version 27 and JASP Version 18. For all tests conducted, my threshold for statistical significance was set at $p < .05$, per the custom in psychological research. When reporting the results of ANOVAs, I report effect sizes using partial eta squared (η_p^2), which is "the proportion of total variation attributable to the factor, partialling out (excluding) other factors from the total nonerror variation." Charles A. Pierce, Richard A. Block & Herman Aguinis, *Cautionary Note on Reporting Eta-Squared Values from Multifactor ANOVA Designs*, 64 EDUC. & PSYCH. MEASUREMENT 916, 918 (2004). Rules of thumb for interpreting partial eta squared suggest a value of .01 reflects a small effect, .06 indicates a medium effect, and .14 reflects a large effect. *See, e.g.*, John T.E. Richardson, *Eta Squared and Partial Eta Squared as Measures of Effect Size in Educational Research*, 6 EDUC. RSCH. REV. 135, 142 (2011).

177. It was evident that participants' judgments could not be explained *purely* by either Hand Formula or Kantian considerations. Combined, economic and Kantian considerations explained about 20.5% of the variance in participants' responses. Participants' judgments were often glaringly at odds with the Hand Formula. For example, the Hand Formula dictates that a defendant need never spend $100,000 (high *B*) to prevent a 20% chance (high *P*) of a $4,000 loss (low *L*) — doing so would be highly wasteful. Yet participants judged Dylan negligent for failing to do this about half (44%) of the time (*EMM* = 4.800). On the other hand, Kantian perspectives cannot fully explain participants' responses either. For example, participants tended to rate the defendant as negligent even when *P* was low (agreeing the defendant was negligent 55% of the time, *EMM* = 4.666). For further discussion of the implications of these findings, see *infra* Section III.A.5.

178. Figure 1 presents a line graph providing a simple two-dimensional visualization of the data from my 2x2x2 design. The use of a line graph is not intended to imply that I used more than two levels ("high" and "low") for any of my experimental manipulations.

FIGURE 1. NEGLIGENCE RATINGS BY *B*, *P*, AND *L* CONDITIONS IN EXPERIMENT ONE[179]



Participants rated the defendant more negligent in conditions where *B* was low and less than *PL* (plotted on the left side of Figure 1[180]) than in conditions where *B* was high and greater than *PL* (plotted on the right side of Figure 1[181]).[182] Participants also rated the defendant more negligent in conditions where *P* was high (represented by the black lines in Figure 1[183]) than in conditions where *P* was low (represented by the gray lines in Figure 1[184]).[185] Participants'

179. The eight dots on the graph reflect the estimated-marginal-mean negligence ratings for the eight possible combinations of the manipulated *B*, *P*, and *L* information, each of which could be either "high" or "low" in the 2x2x2 design.

180. *EMM* = 5.947. I report estimated marginal means (*EMM*s), which "take into account the effects of other variables in the data analysis." *See* Jonathan Cardi, Valerie P. Hans & Gregory Parks, *Do Black Injuries Matter?: Implicit Bias and Jury Decision Making in Tort Cases*, 93 S. CAL. L. REV. 507, 542 n.177 (2020). The *EMM* values do not differ meaningfully from the raw means in my studies.

181. *EMM* = 4.348.

182. $F(1, 206) = 34.863$, $p < .001$, $\eta_p^2 = .145$.

183. *EMM* = 5.629.

184. *EMM* = 4.666.

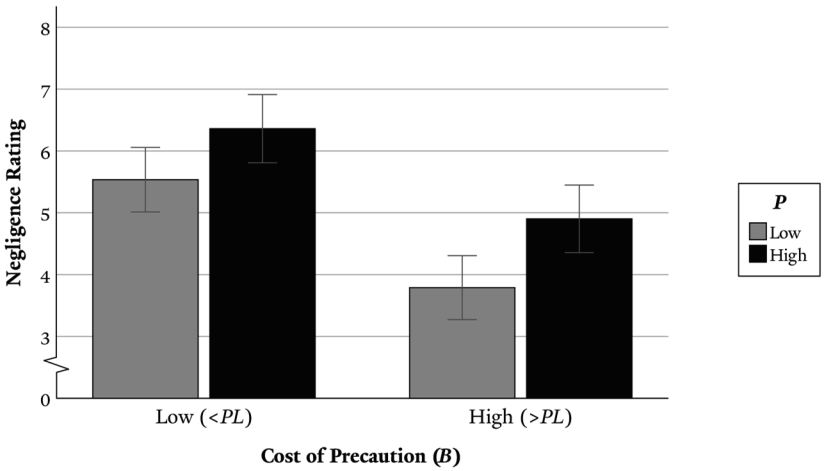
185. $F(1, 206) = 12.658$, $p < .001$, $\eta_p^2 = .058$.

negligence ratings did not differ based on whether *L* was high (solid lines[186]) or low (dotted lines[187]).[188]

As a second way of testing Kantian influence on my participants' judgments, I examined how negligence ratings varied across the three possible values of *PL* (*PL* could equal $8, $800, or $8,000).[189] Though negligence ratings increased as the value of *PL* increased, the effect fell just short of statistical significance, with a p-value of .051.[190] *P* had a more significant influence on participants' judgments in Experiment One than did *L* or the combination of *P* and *L*.

Figure 2 presents the same data as Figure 1 in a streamlined bar graph, depicting the effects of *B* and *P* collapsed across the *L* variable (which had no effect).

FIGURE 2. NEGLIGENCE RATINGS (ESTIMATED MARGINAL MEANS ON EIGHT-POINT SCALE) BY *B* AND *P* CONDITIONS IN EXPERIMENT ONE[191]



186. *EMM* = 5.159.

187. *EMM* = 5.136.

188. F(1, 206) = 0.007, p = .935. Further, there were no statistically significant interactions, with all p's ≥ .118.

189. For values of *PL*, see *infra* Appendix Table 1.

190. For *PL* of $8, *M* = 4.68. For *PL* of $800, *M* = 5.06. For *PL* of $80,000, *M* = 5.69. F(2, 211) = 3.027, p = .051.

191. Error bars reflect 95% confidence intervals.

Finally, after participants judged the mock case (and could no longer change their responses), they were asked to rate the influence that *B*, *P*, and *L* information had on their judgments. Participants reported that *B* information,[192] *P* information,[193] and *L* information[194] affected their judgments roughly equally; participants seemed unaware that *P* bore more weight than *L* in their judgments.[195]

C. *Experiment Two: The Influence of* B, P, *and* L *in a Personal-Injury Case Against an Individual*

The effects of *B* and *P* in Experiment One suggest that economic views and Kantian views capture meaningful influences on lay reasonableness judgments. Experiment Two was designed to test whether those effects would replicate in a different type of tort case.

Participants evaluated a fact pattern loosely inspired by the famous cricket case, *Bolton v. Stone*.[196] This fact pattern differed in that it focused on personal injury, rather than damage to property,[197] and in that risk was spread across multiple unidentified potential victims (parkgoers), rather than focused on one known potential victim (the neighboring farmer).

The effects of *B* and *P* were replicated in Experiment Two.[198] Ultimately, Experiment Two provided further evidence that lay reasonableness judgments are influenced both by economic and Kantian considerations.

1. *The Case*

Danny built a cricket field on land he owned so that his cricket club could play there. When building the field, Danny faced a decision about whether to install safety netting to prevent any cricket balls from flying into the neighboring park. He chose not to install safety netting. When an exceptionally well-hit ball

192. $M = 6.379$.

193. $M = 6.724$.

194. $M = 6.880$.

195. There were no significant differences among the three ratings, $F(2, 426) = 2.628$, $p = .073$ (sphericity assumed per Mauchly's test).

196. [1951] AC 850 (HL) (appeal taken from Eng.).

197. *Cf.* KEATING, *supra* note 71, at 9 (noting that the distinction between injury to property and injury to person has special significance on some accounts of negligence).

198. There was also an effect of *L* in Experiment Two, though I do not dwell on this as it was the only time I observed such an effect.

flew into the park and injured a parkgoer, the parkgoer sued, claiming Danny's failure to install safety netting was negligent.

The vignette again included information about the values for all three Hand Formula inputs, and it noted that Danny was aware of this information when he decided not to install the netting. The *B* paragraph specified the cost of purchasing and installing safety netting: either $4, because the company that built the cricket field had netting left over that they were happy to throw in, or $14,400. The *P* paragraph stated that the probability of a cricket ball striking someone over the next three years in the absence of safety netting—either 1% or 40%—and further explained that the risk was zero with safety netting installed. The *L* paragraph specified that, given the rate at which a cricket ball would have to be moving when it reached the park, any parkgoer it struck would be expected to suffer either minor injuries that would be fully compensated by $600 or very significant, though non-life-threatening, injuries that would be fully compensated by $24,000.[199] The vignette also noted that installing safety netting would not create any other risks or otherwise affect either the cricket field or the public park.

2. *Results*

Participants' negligence ratings[200] were influenced by each of *B* (the cost of the safety netting), *P* (the probability of a stray cricket ball striking a parkgoer absent safety netting), and *L* (how badly a stray cricket ball would be expected to injure a parkgoer).[201] The effects of *P* (and, in this study, *L*) provide further evidence that *PL* information has Hand-independent, Kantian influence on participants' reasonableness judgments.

Figure 3 plots the means of participants' negligence ratings for each combination of *B*, *P*, and *L*.

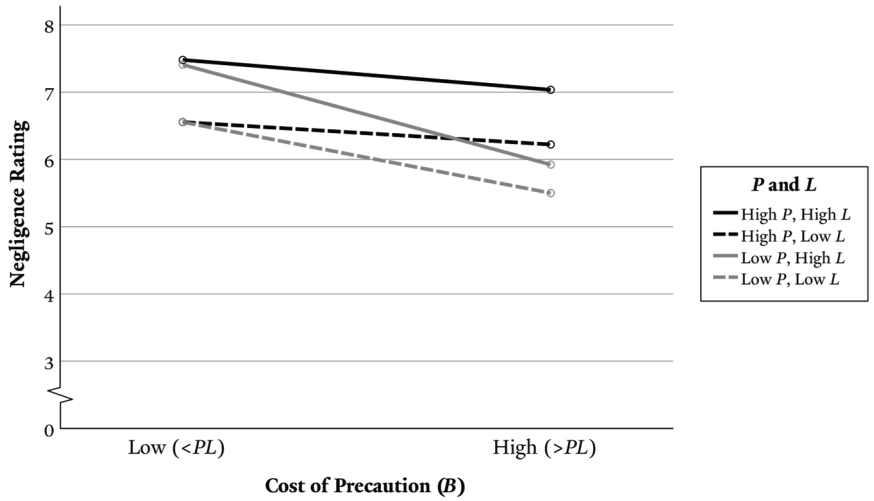
---

199. Thus, as in Experiment One, the magnitude of the *P* and *L* manipulations were equal, and the high *B* value exceeded the highest possible product *PL* by the same proportion that the lowest possible product *PL* exceeded the low *B* value.

200. N = 199. Twenty participants were excluded based on preregistered exclusion criteria.

201. As in Experiment One, I conducted a three-way factorial ANOVA. As with all reported ANOVA tests, the pattern of results is the same with a nonparametric Kruskal-Wallis test. *See supra* note 176.

FIGURE 3. NEGLIGENCE RATINGS BY *B*, *P*, AND *L* CONDITIONS IN EXPERIMENT TWO[202]



As Figure 3 shows, participants rated the defendant more negligent in conditions where *B* was low and less than *PL* (plotted on the left side of the figure[203]) than in conditions where *B* was high and greater than *PL* (plotted on the right side of the figure[204])[205]—consistent with an economic, cost-driven approach. But participants also rated the defendant more negligent in conditions where *P* was high (represented by the black lines in the figure[206]) than in conditions where *P* was low (represented by the gray lines in the figure[207])—reflecting, again, a direct influence of *P* on negligence judgments independent of its role in the Hand Formula.[208] Finally, in this experiment, participants also rated the defendant more negligent when *L* was high (represented by the solid lines in the

202. The eight dots on the graph reflect the estimated-marginal-mean negligence ratings for the eight possible combinations of the manipulated *B*, *P*, and *L* information, each of which could be either "high" or "low" in the 2x2x2 design.

203. *EMM* = 7.000.

204. *EMM* = 6.170.

205. $F(1, 191) = 13.033$, $p < .001$, $\eta_p^2 = .064$.

206. *EMM* = 6.823.

207. *EMM* = 6.347.

208. $F(1, 191) = 4.297$, $p = .040$, $\eta_p^2 = .022$.

figure[209]) than when *L* was low (represented by the dotted lines in the figure[210]).[211] This was the only study in which I observed an effect of *L*.

Paralleling Experiment One, I also examined how negligence ratings varied across the three possible values of *PL* ($6, $240, and $9,600 in Experiment Two). Negligence ratings again increased as the value of *PL* increased, independent of the relation between *PL* and *B*.[212] In this experiment, the effect of *PL* was statistically significant. In fact, *PL* information had a larger effect on participants' negligence ratings than *B* information. Between the effects of *P* alone, of *L* alone, and of *PL*, this experiment provides some evidence that participants were influenced by broad Kantian considerations.[213]

To parallel my presentation of findings in the other reported experiments where *L* had no effect, I include Figure 4, which presents a streamlined bar graph depicting the effects of *B* and *P* collapsed across *L*.

---

209. *EMM* = 6.962.

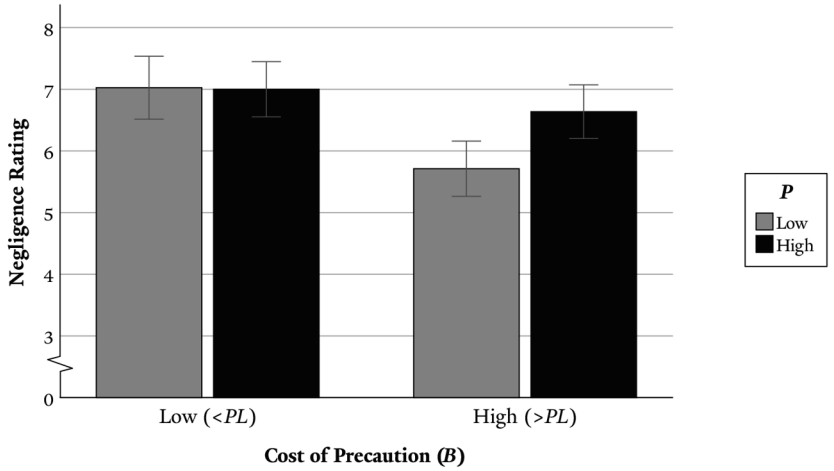
210. *EMM* = 6.208.

211. $F(1, 191) = 10.751$, $p = .001$, $\eta_p^2 = .053$. Visually, Figures 3 and 4 may suggest to some readers a *BP* interaction, with *P* information influencing negligence ratings more when *B* is high than when *B* is low (or, stated differently, with *B* information influencing negligence ratings more when *P* is low than when *P* is high). The *BP* interaction was not statistically significant in this study ($p = .057$), though an interaction of this type was documented in Experiment Five. *See infra* Section II.F.2.

212. For *PL* of $6, $M = 5.93$. For *PL* of $240, $M = 6.49$. For *PL* of $9,600, $M = 7.25$. $F(2, 196) = 7.812$, $p < .001$, $\eta_p^2 = .074$.

213. While both economic and Kantian views appear to capture some aspects of participants' reasonableness judgments, neither fully explained participants' judgments. Participants in Experiment Two were likely to find the defendant negligent even when economic and Kantian views alike indicated he acted reasonably: even when *B* was high, and *P* and *L* were both low, participants found the defendant negligent 69% of the time, giving a mean negligence rating of 5.50. Clearly, there are other considerations influencing participants' judgments. *See infra* Section III.A.5.

FIGURE 4. NEGLIGENCE RATINGS (ESTIMATED MARGINAL MEANS ON EIGHT-POINT SCALE) BY *B* AND *P* CONDITIONS IN EXPERIMENT TWO[214]



Finally, participants self-reported that *B* information,[215] *P* information,[216] and *L* information[217] were equally important to their judgments.[218] In this experiment, all three types of information did, in fact, influence participants' judgments, so there was no detectable disconnect between what participants thought influenced them and what actually influenced them.

D. Experiment Three: The Influence of B, P, *and* L *in a Personal-Injury Case Against a Company*

Experiment Three was designed to investigate whether economic and Kantian influences on participants' reasonableness judgments (observed in both Experiments One and Two) persist when the defendant is an entity rather than an individual. While the outsized Kantian influence of the probability of harm (*P*) persisted, there were no effects of the burdens of precautions (*B*) or the expected severity of harm should harm occur (*L*) in Experiment Three.

---

214. Error bars reflect 95% confidence intervals.

215. $M = 6.467$.

216. $M = 6.648$.

217. $M = 6.598$.

218. $F(2, 396) = 0.269$, $p = .764$ (sphericity assumed per Mauchly's test).

1. *The Case*

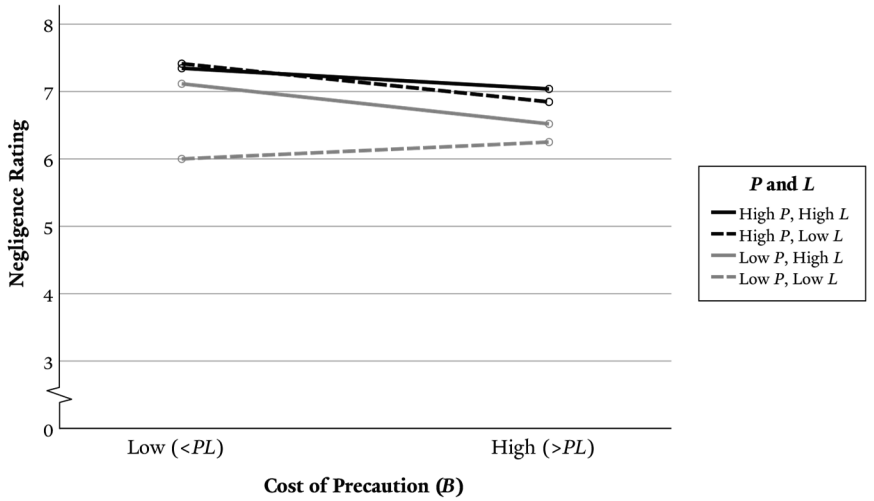
The vignette used in Experiment Three was the same cricket case as that in Experiment Two, but edited to make the defendant an entity that operated a private recreational cricket club rather than an individual cricket enthusiast.

2. *Results*

Participants' negligence ratings[219] were influenced by *P*, but not by *B* or *L*. This provided more evidence that Kantian views capture something important about the lay-legal concept of reasonableness. Indeed, this was the third time in three studies that I observed a significant Kantian effect of *P*.

Figure 5 plots the means of participants' negligence ratings for each combination of *B*, *P*, and *L*.

FIGURE 5. NEGLIGENCE RATINGS BY *B*, *P*, AND *L* CONDITIONS IN EXPERIMENT THREE[220]



219. N = 209. Eleven participants were excluded based on preregistered criteria.

220. The eight dots on the graph reflect the estimated-marginal-mean negligence ratings for the eight possible combinations of the manipulated *B*, *P*, and *L* information, each of which could be either "high" or "low" in the 2x2x2 design.

As Figure 5 shows, participants rated the defendant more negligent in conditions where *P* was high (represented by the black lines in the figure[221]) than in conditions where *P* was low (represented by the gray lines in the figure[222]).[223] Participants' negligence ratings did not vary significantly in response to *B*[224] or *L*[225] in this experiment.[226]

Once again, an alternative analysis focused on the influence of *PL* found that participants rated the defendant significantly more negligent as the value of *PL* increased, independent of the relation between *PL* and *B*.[227] This provided further evidence that participants' negligence ratings were influenced by Kantian considerations.

Figure 6 presents a bar graph reflecting the effect of *P* (and the noneffect of *B*) collapsed across *L* conditions. Many factors might explain *B*'s surprisingly negligible effect in Experiment Three. While participants may have disregarded this factor simply due to chance, laypeople may also be less influenced by the size of the burdens of precaution when they are evaluating the behavior of a corporate defendant, as opposed to an individual defendant.[228]

---

221. *EMM* = 7.161.

222. *EMM* = 6.471.

223. $F(1, 201) = 10.158$, $p = .002$, $\eta_p^2 = .048$.

224. $F(1, 201) = 1.988$, $p = .160$.

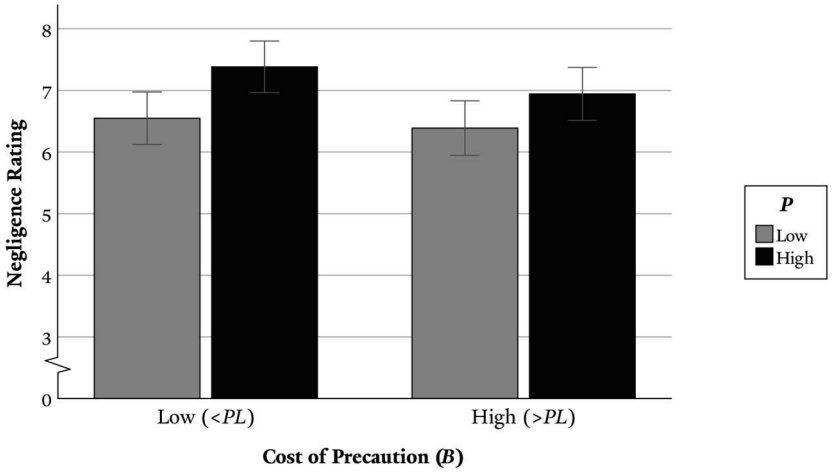
225. $F(1, 201) = 3.042$, $p = .083$.

226. There were no significant interactions (all p's $\geq .147$).

227. For *PL* of \$6, *M* = 6.12. For *PL* of \$240, *M* = 6.99. For *PL* of \$9,600, *M* = 7.19. $F(2, 206) = 7.237$, $p < .001$, $\eta_p^2 = .066$.

228. *See infra* Section III.A.3.

FIGURE 6. NEGLIGENCE RATINGS (ESTIMATED MARGINAL MEANS ON EIGHT-POINT SCALE) BY *B* AND *P* CONDITIONS IN EXPERIMENT THREE[229]



Participants again self-reported that *B* information,[230] *P* information,[231] and *L* information[232] were equally important to their judgments.[233] However, of the three, only *P* information significantly affected participants' judgments in this study.

E. Experiment Four: Probing the Relative Influence of B *and* P

The only independent variable to significantly affect participants' negligence ratings in all of Experiments One, Two, and Three was *P*. These findings support the Kantian view that participants give the risk side of the Hand Formula (particularly, the probability of harm) more weight than the Hand Formula prescribes.

Kantian considerations influence participants' judgments. But do they influence those judgments *more* than economic considerations? At first glance, the reader might think "yes." After all, *P* information had a Kantian influence on participants' judgments in each of the studies examined thus far, while *B*

229. Error bars reflect 95% confidence intervals.

230. $M = 6.789$.

231. $M = 6.732$.

232. $M = 6.684$.

233. $F(2, 416) = 0.091$, $p = .913$ (sphericity assumed per Mauchly's test).

information only influenced these judgments in two out of three studies. But looking at effect sizes, the story is more complicated: in Experiments One and Two, the effect size of *B* was (descriptively, though not statistically) larger than the effect size of *P*. Experiment Four was therefore designed to assess more rigorously whether the (Kantian) influence of *P* on participants' negligence ratings was greater than the (economic) influence of *B*.

Specifically, Experiment Four investigated the possibility that my first three experiments understate the effect size of *P* relative to *B*, because the *P* manipulations were of a smaller magnitude than the *B* manipulations. In Experiment One, the manipulation of the *B* independent variable (with a high value of $100,000 versus a low value of $6) was of a larger magnitude than the manipulation of the *P* variable (with a high value of 20% versus a low value of 0.2%). In Experiments Two and Three, the ratio of high *B* to low *B* (3,600:1) was again far greater than the manipulation of high *P* to low *P* (40:1). Perhaps if the *B* and *P* manipulations were of equal magnitudes, *P* would have a larger effect than *B*. This would provide evidence that participants are not only influenced by Kantian considerations, but also that they are influenced by Kantian considerations *more than* by economic considerations.

To test this possibility, Experiment Four returned to the neighboring-farmers case used in Experiment One, but this time focused only on two independent variables (*B* and *P*), which were both manipulated on a 100:1 scale. I found both a significant (economic) effect of *B* and a significant (Kantian) effect of *P*. The effect size of *P* was descriptively but not statistically larger. *P* again mattered more than the Hand Formula suggests, but I cannot say that Kantian considerations had a statistically greater influence on participants' negligence ratings than economic considerations did.

1. *The Case*

Once again, Dylan's choice of the riskier farming method resulted in a landslide that destroyed his neighbor Pat's barn. The only difference from the neighboring-farmers vignette used in Experiment One was that, in Experiment Four, I assigned different values to *B*, *P*, and *L*.

Specifically, my manipulations in Experiment Four focused on the values of *B* and *P*, using a 2x2 design. The safer farming method, Method A, could be either $90 or $9,000 more expensive than the riskier Method B. And the risk of a landslide could be either 0.6% or 60% with the riskier Method B, as opposed to zero with Method A. Thus, in Experiment Four, the proportion of high *B* to low *B* was the same as the proportion of high *P* to low *P* (both 100 to 1).

I did not manipulate *L* independently of *P* in Experiment Four.[234] When *P* was low, *L* was always assigned the value of $22,500. When *P* was high, *L* was always assigned the value of $10,000. Thus, as in previous experiments, the high *B* value exceeded the highest possible product of *P* and *L* by the same proportion that the lowest possible product of *P* and *L* exceeded the low *B* value. Consequently, as in prior experiments, the economic hypothesis was that participants would rate Dylan more negligent when *B* was low than when *B* was high, and the Kantian hypothesis was that participants would rate Dylan more negligent when *P* (and *PL*) were high than when *P* (and *PL*) were low.[235]

The fact that both *B* and *P* were manipulated on the same 100 to 1 scale in Experiment Four allowed for an apples-to-apples comparison of the effects of *B* and *P* that was not possible in the first three experiments. A larger effect of *B* would provide evidence that people's intuitive reasonableness judgments are more influenced by economic than Kantian considerations; a larger effect of *P* would provide evidence that people's intuitive reasonableness judgments are more influenced by Kantian than economic considerations.

2. *Results*

Once again, participants' negligence ratings[236] were significantly influenced by both the *B* and *P* manipulations.[237] Descriptively, the effect size of *P* was larger than the effect size of *B* in this experiment.[238] The difference in effect sizes was

---

234. Mathematically, with both *B* and *P* manipulated on the same 100:1 scale, I could not manipulate *L* independently while also ensuring that *B* would be less than *PL* when *B* was low and greater than *PL* when *B* was high.

235. The latter hypothesis is not affected by the fact that *L* values always corresponded to *P* values in this study, because *L* was always lower ($10,000) when *P* was high (60%), and *L* was always higher ($22,500) when *P* was low (0.6%). Thus, if participants rate Dylan more negligent when *P* and *PL* are high than when *P* and *PL* are low, that effect could not plausibly be driven by the *L* values more than by the *P* values. Any attempt to explain such an effect as driven by the change in *L* values would entail arguing that people judge others to be significantly less negligent when *L* is high (that is, when the severity of expected harm is great) than when *L* is low, in defiance of theory and common sense. Any effect of the *P* manipulation might reflect either a narrow Kantian effect of *P* or a broad Kantian effect of *PL*, but it could not plausibly reflect an independent effect of *L*. Appendix Table 5, *infra*, provides a detailed summary of my 2x2 design, laying out each possible combination of *B*, on one hand, and *P* and *L*, on the other. It also presents, for each possible combination, what economic and strong Kantian views predict about whether defendant Dylan's actions will be judged negligent.

236. N = 215. Six participants were excluded based on preregistered criteria.

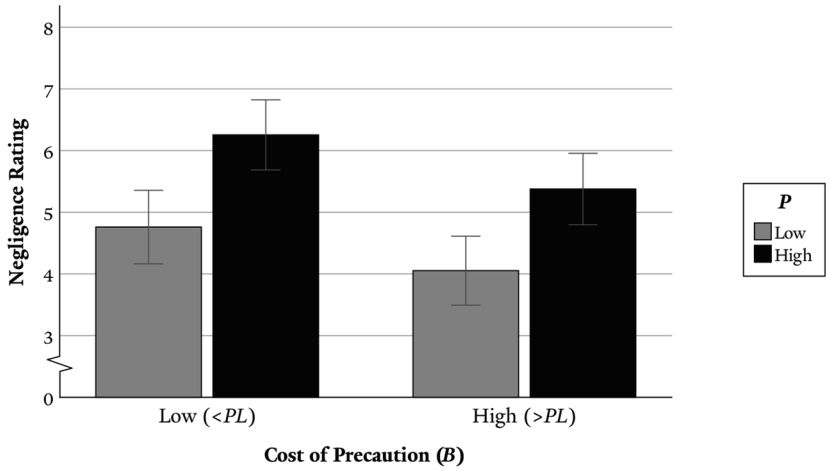
237. There was no need to separately test for an effect of the product *PL* in this experiment: since the *P* manipulation always dictated the *L* value, the effect of *P* and the effect of *PL* are one and the same in this study.

238. $\eta_p^2(P) = .099$; $\eta_p^2(B) = .034$.

not statistically significant, however.[239] Thus, Experiment Four provided further evidence that both the economic and Kantian views capture something significant about lay reasonableness judgments, but did not provide clear evidence that one view captures more about lay reasoning than the other.

Figure 7 shows that participants rated the defendant more negligent when *B* was low and less than *PL*[240] than when *B* was high and greater than *PL*,[241] reflecting some degree of economic influence on their judgments.[242] But participants also rated the defendant more negligent when *P* was high[243] than when *P* was low,[244] reflecting a distinct Kantian influence of *P* on their judgments.[245]

FIGURE 7. NEGLIGENCE RATINGS (ESTIMATED MARGINAL MEANS ON EIGHT-POINT SCALE) BY *B* AND *P* CONDITIONS IN EXPERIMENT FOUR[246]



239. A test of the linear combination of the parameter estimates for the effects of *B* and *P* did not reject the possibility that the two effects were equal.

240. *EMM* = 5.507.

241. *EMM* = 4.715.

242. $F(1, 211) = 7.361$, $p = .007$, $\eta_p^2 = .034$.

243. *EMM* = 5.816.

244. *EMM* = 4.406.

245. $F(1, 211) = 23.302$ (sphericity assumed per Mauchly's test), $p < .001$, $\eta_p^2 = .099$. Again, there was no interaction ($p = .772$).

246. Error bars reflect 95% confidence intervals.

Notably, in Experiment Four, participants' self-assessment of what influenced their reasonableness judgments did not align with their actual judgments.[247] Participants rated *L* information as most influential[248] and significantly more influential than *B* information,[249] which was rated least influential.[250] Yet, given the manipulations in this study, participants actually rated Dylan *more* negligent when *L* was low than when *L* was high, while the manipulation of *B* had a significant effect in the expected direction. These results provide further evidence that participants are not fully cognizant of how *B*, *P*, and *L* information influences their judgments of whether conduct is reasonable.

F. Experiment Five: The Persistence of Kantian Influence

Experiments One through Four paint a consistent picture: *PL*, and particularly *P*, gets more weight in lay reasonableness judgments than the Hand Formula suggests. But a skeptical reader may wonder whether the consistent effects in these studies truly reflect something Kantian in participants' decision-making, or whether they instead reflect participants' laziness in the pursuit of cost-benefit analysis. Specifically, a critic might posit that participants are generally inclined to apply the Hand Formula to assess reasonableness but are either unable or unwilling to do the mathematical work needed to compare *B* to *PL* in the context of my experiments, and therefore simply look to *P* as a mental shortcut or heuristic. Experiment Five addresses this potential critique by replicating Experiment Four, but with the math done — and the cost-benefit conclusion plainly stated — for participants.

1. The Case

The case vignette used in Experiment Five was the same as that in Experiment Four, but with the *BPL* math laid out for participants. Specifically, a new paragraph was added to the vignette after the *B*, *P*, and *L* paragraphs. This paragraph explained that using Method A rather than Method B would cost Dylan more or less (depending on the condition) than it would be expected to save in landslide-related harms. The paragraph ended with a direct statement that "cost-

247. Participants differentiated among the influences of *B*, *P*, and *L* in this experiment. $F(2, 428) = 3.096$, $p = .046$, $\eta_p^2 = .014$.

248. $M = 6.767$.

249. $M = 6.200$.

250. Mean difference = 0.567, $p = .025$.

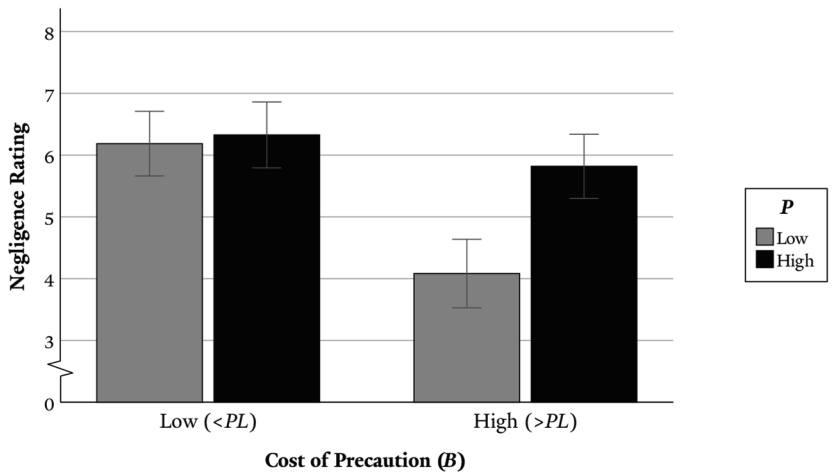
benefit analysis therefore suggests that Dylan should use" Method A or Method B, depending on the experimental condition.[251]

Relatedly, when participants were asked to self-report how much of a role the *B*, *P*, and *L* information played in their judgments, they were also asked to rate the role of the fact that "cost-benefit analysis suggests Dylan should use" Method A or Method B.

2. *Results*

Experiment Five replicated Experiment Four: participants' negligence ratings[252] were again influenced by both the *B* and *P* manipulations. Importantly, this means that *P* continued to have an outsized influence on participants' judgments with the Hand Formula math laid out. This provides strong support for the Kantian hypothesis that the significance of the risk factors into lay reasonableness judgments more than suggested by the Hand Formula.

FIGURE 8. NEGLIGENCE RATINGS (ESTIMATED MARGINAL MEANS ON EIGHT-POINT SCALE) BY *B* AND *P* CONDITIONS IN EXPERIMENT FIVE[253]



251. The cost-benefit conclusion was reemphasized in a bullet point that appeared in the bolded summary paragraph after the bullet-point summaries of the *B*, *P*, and *L* information. Given that the order in which the *B*, *P*, and *L* information was presented had not significantly affected participants' negligence ratings in any prior experiments, all Experiment Five participants viewed the materials in *B*-*P*-*L* order for ease of readability. *See supra* note 157.

252. N = 209. Twelve participants were excluded based on preregistered criteria.

253. Error bars reflect 95% confidence intervals.

Figure 8 shows that participants again rated the defendant more negligent when *B* was low and less than *PL*[254] than when *B* was high and greater than *PL*.[255] Participants also rated the defendant more negligent when *P* was high[256] than when *P* was low.[257] In this experiment, there was a significant interaction between *B* and *P*.[258]

The effect size of *B* was descriptively, though not statistically, larger than the effect size of *P* in this experiment.[259] This could reflect that participants were more concerned with cost justification than probability of harm, but it could also simply reflect a demand characteristic of the study. That is, spelling out cost-benefit analysis for the participants and telling them whether cost-benefit analysis indicated Dylan should use Method A or Method B may have led participants to believe that they were supposed to base their reasonableness judgments on these factors, leading them to place more weight on cost-benefit considerations.

Participants' self-reported ratings of how the different pieces of information influenced their negligence judgments showed far greater variation in this experiment than in the first four. This may be due to the inclusion of the item asking them to rate the influence of the statement that cost-benefit analysis suggested Dylan should use Method A or Method B (as applicable). Participants rated *P* information as most influential in this study,[260] followed by (in descending order) *L* information,[261] the cost-benefit statement,[262] and finally *B*

---

254. *EMM* = 6.256.

255. *EMM* = 4.951. $F(1, 205) = 23.350$, $p < .001$, $\eta_p^2 = .102$.

256. *EMM* = 6.073.

257. *EMM* = 5.134. $F(1, 205) = 12.066$, $p < .001$, $\eta_p^2 = .056$.

258. $F(1, 205) = 8.696$, $p = .004$, $\eta_p^2 = .041$. This interaction effect can be stated either of two ways: the effect of *P* was greater in high *B* conditions than in low *B* conditions, or the effect of *B* was greater in low *P* conditions than in high *P* conditions. In short, conditions that included either a low *B* value or a high *P* value pushed participants to rate the defendant more negligent (all means ≥ 5.82), while defendants were rated less negligent when *B* was high and *P* was low ($M = 4.08$). This pattern of findings has interesting parallels with Lord Reid's post-*Bolton* suggestion, in *Wagon Mound No. 2*, that while the burden of precautions is not relevant in cases involving real and significant risks, it is relevant where a case involves real but insignificant risks, "because an actor who can easily eliminate even such risks should not impose them." Weinrib, *supra* note 78, at 519 n.61 (citing Overseas Tankship (U.K.) Ltd. v. Miller Steamship Co. (*Wagon Mound No. 2*) [1967] AC 617 (PC) 641 (appeal taken from N.S.W.)).

259. $\eta_p^2(B) = .102$; $\eta_p^2(P) = .056$. A test of the linear combination of the parameter estimates for the effects of *B* and *P* did not reject the possibility that the two effects' sizes were equal.

260. $M = 7.742$.

261. $M = 6.880$.

262. $M = 6.301$.

information.[263] Paired comparisons showed that participants' ratings of each piece of information differed significantly from their ratings of every other piece of information.[264] It is particularly interesting that, when given the chance to explicitly state whether *P* information or cost justification was more important to them, participants rated the *P* information as more important.

III. REALIGNING THE HAND FORMULA

The Hand Formula is out of alignment with how laypeople judge whether conduct is reasonable. My experiments indicate that though laypeople are influenced to a degree by the balance between the burden of precautions and the risks those precautions would mitigate, the Hand Formula fails to capture the outsized Kantian weight that they place on the risk side of the ledger.

As Table 2 shows, *PL*—and particularly *P*—had a consistent influence beyond that posited by the Hand Formula. In four of five studies, participants rated the defendant more negligent as the value of the product *PL* increased, independent of whether *B* was greater or less than *PL* (the one exception being the p = .051 in Experiment One). And in all studies, participants rated the defendant more negligent when *P* was high than when *P* was low, independent of whether *B* was greater or less than *PL*. These findings indicate that participants consistently evaluated reasonableness in partially Kantian terms—that is, the significance of the risk the defendant created influenced participants' judgments independent of Hand balancing.

TABLE 2. SUMMARY OF EFFECTS OF *B*, *P*, AND *L* INDIVIDUALLY, AND OF THE PRODUCT *PL*, ON NEGLIGENCE RATINGS ACROSS EXPERIMENTS[265]

	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
B* < *PL* or *B* > *PL	⋆	⋆		⋆	⋆
P	⋆	⋆	⋆	⋆	⋆
L		⋆			
PL	p = .051	⋆	⋆	⋆	⋆

263. M = 5.177. $F(2.536, 527.394) = 40.425$, $p < .001$, $\eta_p^2 = .163$. For Experiment Five, Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(5) = 57.063$, $p < .001$), so degrees of freedom were corrected using the Greenhouse-Geisser correction, $\epsilon = .845$.

264. All p's ≤ .017.

265. ⋆ signifies that the variable had a significant effect. Grayed cells were not tested.

In discussing my findings, I have focused on the ways that *B*, *P*, and *L* information affected participants' negligence ratings on an eight-point scale. One might reasonably wonder whether *B*, *P*, and *L* information had similar effects on participants' binary *verdicts* of "negligent" versus "not negligent." In short, they did.[266]

Participants' responses, particularly their brief written explanations of their negligence ratings, suggest significant heterogeneity in how they think about reasonableness; some participants express more economic views and others more Kantian. It is quite possible that certain individuals would consistently decide cases using either economic or Kantian reasonableness criteria. My analyses, by their nature, focused on what influenced participants' judgments at the aggregate level. My findings indicate significant Kantian influences across a large sample of participants making lay-legal reasonableness judgments.

In the remainder of the Article, I explain some of the implications of these findings for tort theory and practice, review limitations of my experimental methodology, and identify promising directions for future empirical research.

A. Discussion of Key Findings

1. The Kantian Effect

The Article's headline finding is the consistent Kantian effect of *PL*—particularly *P*—on participants' reasonableness judgments. My studies provide the first experimental evidence that Kantian views of reasonableness capture something significant about lay reasonableness judgments. When laypeople decide whether a tort defendant acted unreasonably, they are influenced by the

266. My use of an eight-point scale with no "neutral" option allowed me to recode participants' negligence ratings as binary verdicts of "not negligent" (negligence ratings one through four) or "negligent" (negligence ratings five through eight). I then conducted post hoc, generalized linear model analyses to examine whether *B*, *P*, and *L* information influenced participants' binary verdicts. I found essentially the same pattern of results presented in Table 2, with only small differences pertaining to Experiments Two and Three. In Experiment Two, the effects of *B* ($p = .055$) and *P* ($p = .147$) on binary verdicts fell short of significance, though both significantly affected eight-point negligence ratings. In Experiment Three, *L* information significantly affected binary verdicts ($p = .021$), though it did not affect eight-point negligence ratings. Overall, analyses of binary verdicts revealed that the Kantian effects in my studies do not just nudge negligence ratings from five to seven or six to eight; *PL*, and particularly *P*, also significantly influenced whether participants deemed the defendant negligent or nonnegligent. The OSF page contains a summary table of my analyses with binary verdicts, together with SPSS output from the generalized linear model analyses. *See* Christopher Brett Jaeger, *The Hand Formula's Unequal Inputs — Generalized Linear Model Analyses with Binary Negligence Verdicts as Outcome Variable*, OSF (July 24, 2025), https://osf.io/9ha7p?view_only=bb29d8f003854756a4f1b84c3ed7d5a7 [https://perma.cc/QL93-JJNP].

significance of the risk the defendant's actions created, independent of whether that risk was cost-justified.[267] Specifically, participants judged defendants more negligent as the value of *PL* increased, whether *PL* was greater than *B* or not. Indeed, the value of *PL*, by itself, explained more variance than Hand cost justification in a majority of experiments, suggesting that Kantian considerations explain as much about lay judgments as economic considerations do.[268]

Identifying precisely which subtype of Kantian theory my results favor is more challenging. The most consistent effect, present in all five studies, was the (Hand-independent) effect of the probability of harm (*P*). But, taking a step back, four of five studies indicated that burdens (*B*) still had some role in participants' evaluations of reasonableness. While the probability of harm (*P*), or the probability of harm times the expected severity of harm (*PL*), got more weight relative to burden (*B*) than the Hand Formula suggests, these results can plausibly be interpreted as supporting either strong or softer Kantian perspectives.[269] On one hand, laypeople might process both the balance of burden and risk reduction *and* the significance of risks in a strong Kantian sense, with both processes contributing to their ultimate judgments. But on the other hand, participants might have processed *B*, *P*, and *L* together but discounted *B* — putting a thumb on the *PL* side of the scale — consistent with what I call softer Kantian views.[270]

Future work might endeavor to tease apart these possibilities.[271] Presently, though, the key takeaway is that lay reasonableness judgments are influenced by Kantian consideration of the significance of risks in ways the Hand Formula does not capture. While more work is needed before we can confidently generalize from my vignette studies to real-world courtrooms,[272] my findings provide support for the notion that Kantian views capture a meaningful component of lay reasonableness judgments.

---

267. *See* Wright, *supra* note 5, at 256 (outlining a Kantian view of reasonableness).

268. Descriptively, *PL* explained more variance than *B* in Experiments Two, Three, and Four. *See supra* Sections II.C, II.D, II.E. In Experiment Four, *P* and *L* were not independently manipulated, so the effect size of *P* was also the effect size of *PL*.

269. *See supra* Table 1.

270. *See supra* text accompanying notes 86-87.

271. Such work could identify and test specific circumstances where strong Kantian views and softer Kantian views yield competing predictions about whether conduct is reasonable.

272. *See infra* Section III.C.

2. *The Primacy of Probability of Harm*

My experiments suggest *P* plays a bigger role than *L* in lay negligence judgments. In all experiments, participants rated the defendant more negligent for failing to take precautions against high-*P* risks than against low-*P* risks (independent of whether those risks were cost-justified). In contrast, *L* only had a significant Hand-independent influence on negligence ratings in Experiment Two. These findings accord with narrow Kantian views of reasonableness. But it is worth noting that prospect theory — prominent in the field of behavioral economics — also suggests that *P* should be more influential than *L*.

Prospect theory involves decision-making under uncertainty.[273] One of the foundational tenets of prospect theory is that losses loom larger than gains: the pain people experience from losing $100 is greater than the joy they experience from gaining $100.[274] Thus, when given the choice, people often seek to maximize their chances of avoiding a loss, preferring to take smaller chances of bigger losses over bigger chances of smaller losses.[275] For example, given a choice between a 45% chance of losing $6,000 and a 90% chance of losing $3,000, people overwhelmingly opt for the former; the more probable loss of $3,000 is viewed as worse, even though the expected value of these two options is identical.[276] One way to frame this finding is that, in the context of risky losses, people give more weight to the probability than to the size of the potential loss.

In some sense, negligence cases ask jurors to evaluate risky losses that were faced by the plaintiff. From a Kantian perspective, jurors in negligence cases judge whether the defendant imposed "significant" unaccepted risks of loss on the plaintiff.[277] Just as people typically consider a higher-probability risk of a smaller loss to be worse than a lower-probability risk of a larger loss, they may be similarly biased to consider a tort defendant's imposition of a higher-probability risk of a smaller loss to be worse than a lower-probability risk of a larger

273. *See* Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision Under Risk*, 47 ECONOMETRICA 263, 263 (1979).

274. *Id.* at 279.

275. *Id.* at 268 (observing that people are "risk seeking in the negative domain"). There is, however, evidence that the tendency toward risk seeking in the loss domain may depend on the probability associated with losses, Amos Tversky & Daniel Kahneman, *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, 5 J. RISK & UNCERTAINTY 297, 297-99 (1992), or on the magnitude of the losses being risked, Marc Scholten & Daniel Read, *Prospect Theory and the "Forgotten" Fourfold Pattern of Risk Preferences*, 48 J. RISK & UNCERTAINTY 67, 69-70 (2014).

276. *See* Kahneman & Tversky, *supra* note 273, at 267-68 (summarizing these findings).

277. *See* Wright, *supra* note 5, at 256.

loss. In Hand Formula terms, this would mean that fact finders place more weight on *P* than *L*, as my participants tended to do.[278]

Thus, the primacy of *P* makes good sense from the perspective of Kantian legal theory, but it also fits well with what psychologists and behavioral economists understand about human judgment and decision-making. In this way, my studies suggest that a revision to the economic assumptions underlying the Hand Formula may be needed to better capture human behavior, echoing prospect theory's revisions to the assumptions of expected-utility theory. My findings suggest that in the behavioral Hand Formula, *P* has an outsized influence.

3. Understanding the Effect of Burdens

In four of five studies, manipulating *B* caused participants to rate defendants as more negligent when *B* was low and less than *PL* than when *B* was high and greater than *PL*. Interpreting my findings in light of economic theory suggests that participants' negligence ratings were influenced by whether the untaken precaution was cost-justified under the Hand Formula.

But it is worth noting a potential alternative explanation of at least some of the influence of the *B* manipulation in my studies. Recall that, throughout my studies, I assigned values of *B* such that the low *B* value was always less than *PL* (meaning the defendant had failed to take a cost-justified precaution), and the high *B* value was always greater than *PL* (meaning the defendant had not failed to take a cost-justified precaution). Thus, the Hand Formula would always indicate negligence when *B* was low and nonnegligence when *B* was high. This approach to assigning values has an enormous virtue: it allows me to confidently infer that the effects of *P*, *L*, and *PL* observed in my experiments reflect influence independent of their relation to *B* under the Hand Formula. But the tradeoff of this approach is the possibility that the observed effects of *B* are attributable not only to its rational effects under the Hand Formula, but also to any irrational influence of *B* beyond what the Hand Formula predicts.

In other words, it is possible that some of the observed influence of *B* is a product of some heuristic based on the *B* value, rather than Hand-style cost-benefit analysis.[279] A *B*-driven heuristic would be something like the following: defendants are liable when it is cheap and easy to be safer (low *B*), but

---

278. Analogously, in criminology, evidence demonstrates that "certainty of apprehension [akin to *P*] is a more effective deterrent than the severity of the ensuing legal consequences [akin to *L*]." Daniel S. Nagin, *Deterrence in the Twenty-First Century*, 42 CRIME & JUST. 199, 215 (2013).

279. While participants using *B* as a heuristic in their judgments would, by definition, mean that they give some weight to the cost of precaution (which would be inconsistent with strong Kantian views), they would not be using *B* in the rational way prescribed by the Hand Formula.

defendants are not liable when it is hard and expensive to be safer (high *B*), *regardless* of how *B* compares to *PL*. Such a heuristic could be especially powerful when *B* is very low, to the point of de minimis, as it was in Experiments One through Three. I am not aware of any theory of reasonableness in the legal scholarship that suggests *B* has or should have an influence independent of its relation to *PL*, but it is certainly within the realm of empirical possibility.

Thus, the effects of *B* in my studies may be best thought of as an upper bound on the influence of cost justification under the Hand Formula. Given my manipulations, the effect of *B* should fully capture any economically rational influence of cost justification, but it may also capture irrational, heuristic influences independent of the variable's role in the Hand Formula. Future research might seek to tease the two apart. I offer some thoughts on how researchers could explore this question in Section III.D below.

Even if the observed effects of *B* on participants' judgments were entirely "rational"—driven by the relationship between *B* and *PL* rather than any heuristic—this still would not suggest that Hand-style cost-benefit analysis played a *leading* role in participants' judgments. Even where I observed an effect of *B*, participants' judgments as to the defendant's negligence still tended to run contrary to the Hand Formula: in four of five experiments, participants tended to find the defendant negligent even when $B > PL$. I discuss this point further in Section III.A.5, *infra*.

Finally, the fact that I observed effects of *B* in four of five studies naturally invites consideration of the one noneffect (Experiment Three). One explanation is, as always, chance. But it is perhaps noteworthy that Experiment Three was the only experiment involving a defendant entity. Prior research suggests that lay decision-makers hold corporate entities to a higher standard of care than individual defendants.[280] Perhaps one mechanism for this differential treatment is that people are less influenced by high (versus low) values of *B* when the defendant is an entity than when the defendant is an individual. Put differently, people may be more willing to excuse individuals for failing to take difficult or expensive precautions than they are to excuse corporate entities for failing to take similar measures.[281] Future research might explore whether people consistently

280. *See* Valerie P. Hans & M. David Ermann, *Responses to Corporate Versus Individual Wrongdoing*, 13 LAW & HUM. BEHAV. 151, 162 (1989) ("Even when their actions are identical, corporations and individuals are judged in divergent ways. Our respondents held the corporation to a higher standard, judged it to be more reckless and more morally wrong in its behaviors, and punished it more severely.").

281. *See, e.g.*, Francis H. Bohlen, *Mixed Questions of Law and Fact*, 72 U. PA. L. REV. 111, 119 (1924) (suggesting that a "corporation [defendant] . . . is always assumed to have unlimited resources").

assign *B* less weight in the calculus of negligence when the defendant is a corporate entity.

4. The Potential for Influence Outside of Awareness

Across all studies, *P* demonstrated an outsized weight on participants' judgments. Yet when self-reporting how much *B*, *P*, and *L* information influenced their judgments, participants only rated *P* as significantly more influential than *L* in one study (Experiment Five). This illustrates that participants' self-reported descriptions of their reasoning did not necessarily reflect what actually drove their judgments. While this is not particularly surprising from a psychological perspective, it nevertheless highlights important limitations on other methodologies examining the nature of reasonableness, such as introspection or surveys asking laypeople to define reasonableness or what they believe would influence their reasonableness judgments. Experiments that directly ask participants to judge reasonableness (like those in this Article) offer a way around these limitations.

5. The Big Picture

This Article has focused on what *B*, *P*, and *L* information can explain about lay negligence judgments, but it is also worth considering what *B*, *P*, and *L* leave unexplained. Economic and Kantian perspectives alike are intended to provide complete accounts of what separates reasonable conduct from unreasonable conduct. These theories suggest that negligence depends entirely on whether the defendant failed to take a cost-justified precaution or whether the defendant imposed significant risks on nonconsenting others. Yet across five studies—each with large manipulations of *B*, *P*, and *L*—the economic and Kantian views never combined to explain more than 21% of the variance in participants' negligence judgments.[282] In the world of social science, that figure is high, yet the vast majority of this variance is still left unexplained.[283]

Consider that, in Experiments Two and Three, participants overwhelmingly found negligence even when *B* was high and *P* and *L* were both low—a situation where economic and Kantian views alike suggest the defendant acted reasonably. Though economic and Kantian considerations influenced participants to a degree, their ultimate negligence judgments were contrary to the Hand Formula

---

282. The highest percentage of variance explained by the overall model (including *B*, *P*, *L*, and interactions) in any of my experiments was 20.5%, in Experiment One.

283. *See* Richardson, *supra* note 176, at 142 (describing η_p^2 of .14 as a large effect size).

and to Kantian theory. This begs the question: what other factors influenced their judgments?

This question provides ample fodder for future research. Tort scholarship offers many additional theories of reasonableness that might help to explain some of the variance left unexplained by *B*, *P*, and *L*.[284] For example, my own earlier work suggests that whether the defendant's conduct was customary does more to influence lay negligence judgments than cost justification.[285]

Research has also uncovered evidence of influences that have less to do with legal theory and doctrine than with human decision-making and biases.[286] For instance, while "'black letter' negligence doctrine states unequivocally that breach can and should be assessed by looking only at the defendant's outward conduct"—and not the defendant's subjective mental state—recent empirical work suggests that jurors are more likely to find tort defendants blameworthy and liable when those defendants have subjective knowledge of the risks their actions will create for others, at least when their subjective belief is that the risk is high.[287] This could help to explain the general tendency of participants in my studies to find the defendant negligent, as defendants in my vignette were conscious of the risks. Another possibility lurking in the background is that even when some people are ostensibly evaluating reasonableness, they may be implicitly applying what is, in effect, a strict-liability standard.[288]

While there are many potential influences on reasonableness judgments to sort through, research testing which of these factors have real explanatory power—and how much—can help illuminate the criteria of reasonableness as they are applied in the tort system.[289]

---

284. For examples, see *supra* note 1.

285. Jaeger, *supra* note 140, at 948.

286. *See supra* note 125 and accompanying text.

287. Margoni & Brown, *supra* note 125, at 1 (emphasis omitted).

288. *See* Hetcher, *supra* note 114, at 643-44; Joseph Sanders, Matthew B. Kugler, Lawrence M. Solan & John M. Darley, *Must Torts Be Wrongs? An Empirical Perspective*, 49 WAKE FOREST L. REV. 1, 2 (2014).

289. There are, of course, myriad ways that tort cases can vary that might affect the criteria applied to negligence cases. For example, the criteria of negligence might vary based on whether cases involve commissions and omissions, or based on whether the alleged tortfeasor was acting in pursuit of private benefit or public benefit when the plaintiff was injured. *See* Wright, *supra* note 5, at 261-74 (cataloguing various types of cases for which the criteria of negligence might vary); Alex Stein, *The Domain of Torts*, 117 COLUM. L. REV. 535, 541 (2017) ("Whether a court deciding a tort case should apply the welfare-oriented rules rather than the fairness-based set of rules, or vice versa, depends on the private or public categorization of the benefit generated by the actor's risky activity."). Future research might also examine the extent to which jury instructions influence lay reasonableness criteria. *See* Jaeger, *supra* note 140, at 944.

B. The Implications of a Misaligned Hand Formula

The debate surrounding tort law's reasonableness standard has both descriptive and normative dimensions. This Article's contributions are primarily descriptive. Law largely leaves the question of whether the defendant acted with reasonable care to lay jurors, and this Article sheds light on how laypeople answer that question. My findings suggest economic and Kantian considerations play meaningful roles. Generalizing from my studies to real-world juror decision-making comes with all the caveats about external validity that necessarily accompany experimental legal research based on written vignettes.[290] More work is certainly needed to jump confidently from the lab to the courtroom. That said, if real jurors deciding whether a tort defendant acted with reasonable care are influenced by the same sorts of considerations that my participants were, then evidence about the "significance" of risks—and, specifically, evidence about the probability of the defendant's conduct causing harm—would play a larger role in their judgments than the Hand Formula predicts.

Sharpening our descriptive account of the criteria of reasonableness is itself a meaningful contribution to the jurisprudential debate.[291] But these arguments likely have meaningful implications for the normative question of how law should define "reasonable care." To be sure, philosophers, scientists, and law professors are generally and rightfully skeptical about reasoning from "is" to "ought." Yet, "[i]n evaluating some normative questions about law . . . it is beneficial to elucidate the relevant ordinary notions."[292] When the criteria that courts and legal scholars use to define a concept vary meaningfully from the criteria that laypeople use, it is at least "worthwhile to evaluate whether the differences in criteria are desirable."[293]

This is especially true with respect to tort reasonableness. Again, tort law traditionally delegates the reasonableness question to the jury with minimal instruction.[294] Thus, jurors have unusual power to dictate what tort law requires

---

290. *See, e.g.*, J. Shahar Dillbary, Cherie Metcalf & Brock Stoddard, *Incentivized Torts: An Empirical Analysis*, 115 NW. U. L. REV. 1337, 1390 n.158 (2021) (discussing external validity).

291. *See* Alan Calnan, *The Nature of Reasonableness*, 105 CORN. L. REV. ONLINE 81, 81 (2020) (arguing that "greater clarity can be gained by taking a scientific approach to [reasonableness], exposing the natural foundations beneath the concept's varied interpretations").

292. Tobia, *supra* note 124, at 2508. Indeed, at times, some scholars have argued that lay understanding ought to have normative force—that law, or certain areas of law, should bend toward the understanding of the people it governs. *See, e.g.*, Joshua Kleinfeld, *Manifesto of Democratic Criminal Justice*, 111 NW. U. L. REV. 1367, 1405 (2017).

293. Jaeger, *supra* note 125, at 437.

294. *See supra* notes 25-28 and accompanying text.

of the people in their community.[295] Some tort theorists have argued—and I agree—that this reflects a democratic deference to community mores; that is, the structure of tort law itself gives some credence to the idea that tort law's reasonable-care standard should reflect the views and beliefs of laypeople in the community.[296] "From the perspective of democratic theory, the jury norm effect's impact on formal legal outcomes is an anti-elitist, liberal feature of American tort law that distinguishes it from its counterparts abroad."[297] Catharine Pierce Wells has argued forcefully that the process of jury decision-making in tort negligence cases, in which each juror brings their own experiences and viewpoints to bear in working toward consensus, endows negligence jury verdicts with the normatively desirable quality of "local objectivity"—a quality that resolutions of negligence claims by a judge lack.[298] Further, theorists have argued that it is desirable for a legal system to use "ordinary concepts accessible to laypeople."[299] If this is the case, one could argue that it is desirable for the "considerations [that] inform judgments of reasonableness [to] be determined by laypeople."[300]

Each of these notions—democratic delegation of the reasonable-care issue to juries, the local objectivity of jury negligence decisions, and the idea that law is better for employing accessible concepts—tends to suggest that alignment between the judicial concept of reasonableness articulated and applied by judges and the lay-legal concept of reasonableness held by the laypeople who may serve

---

295. *See* Kelley, *supra* note 3, at 352 (noting the primacy of the jury in applying the reasonable-person standard).

296. *See* Hetcher, *supra* note 114, at 654 ("[T]he jury's role in American tort litigation is normatively significant in a number of respects, both substantive and procedural."); Tilley, *supra* note 2, at 1324 (contending that "tort law operates as a vehicle through which communities perpetually reexamine and communicate their values"); *see also* Martin A. Kotler, *Social Norms and Judicial Rulemaking: Commitment to Political Process and the Basis of Tort Law*, 49 KAN. L. REV. 65, 98-99 (2000). For an opposing view, see VALENTIN JEUTNER, THE REASONABLE PERSON: A LEGAL BIOGRAPHY 68 (2024), which argues that the common law's reasonable-person standard was "meant to be . . . a standard controlled by the courts."

297. Hetcher, *supra* note 114, at 636; *see also* Abraham, *supra* note 28, at 1198-99 (expressing skepticism that Americans would tolerate "having individual judges assess the quality of everyday conduct in order to decide whether to impose civil liability for the consequences of that conduct" and that leaving that issue to the jury is "a testament to the growing influence of the democratic ethos as negligence law has developed").

298. Catharine Pierce Wells, *Tort Law as Corrective Justice: A Pragmatic Justification for Jury Adjudication*, 88 MICH. L. REV. 2348, 2409-10 (1990) ("[A] judgment is locally objective if made by consensus after a process of investigation and deliberation that, as a matter of practical experience, produces agreements in local normative judgments among persons with varying normative viewpoints.").

299. Martínez & Tobia, *supra* note 70, at 181-82.

300. *Id.* at 182.

as jurors is normatively desirable.[301] My empirical findings suggest that the Hand Formula suffers from a misalignment in this respect. These results suggest that the Hand Formula, standing alone, fails to capture important aspects of the lay-legal concept of reasonableness. This misalignment is arguably a normative strike against the Hand Formula. It casts doubt on the proposition that the Hand Formula, by itself, provides the best conception of tort reasonableness, as other conceptions (e.g., hybrid standards of reasonableness[302]) align better with lay understanding.

To illustrate the problems that this sort of misalignment poses, imagine a real-world negligence case in which complete *B*, *P*, and *L* information happens to be available and undisputed. Assume (consistent with my empirical findings) that jurors in the relevant community care to some degree about whether the defendant's conduct was cost-justified, but that cost justification is not the primary driver of their negligence judgments. Now, imagine that a court in the same community decides a negligence case at the summary-judgment stage, solely on Hand Formula grounds. In taking the issue from the jury and resolving it with different criteria, the court undermines the democratic values reflected in tort law's delegation of negligence judgments to lay understanding, the "local objectivity" baked into the process of jury decision-making, *and* the desirable legal feature of conceptual accessibility (as the concept of reasonableness deployed by the court differs from that held by the community).[303] The potential for such disconnects between judicial and lay understandings suggests that courts should be wary of taking negligence disputes away from juries.[304]

---

301. *See supra* Section I.C (distinguishing lay, lay-legal, and legal concepts of reasonableness).

302. Tobia, *supra* note 1, at 296 (asserting that reasonableness is best understood as a "hybrid notion that is partly statistical and partly prescriptive").

303. Jaeger, *supra* note 125, at 438.

304. What if the negligence case we are imagining was always headed for a *bench trial* rather than a jury trial? When deciding whether the defendant acted unreasonably, should the court (try to) conform its reasoning to the criteria that would be used by lay jurors in the community (or that the court believes would be used by lay jurors in the community)? For example, if the relevant community holds particularly Kantian views about reasonable care, should the court eschew its preferred Hand Formula to try to evaluate the case using Kantian criteria?

On the one hand, a normative focus on maximizing the alignment of the legal and lay-legal concepts of reasonable care—together with a concern that like cases be treated alike—would indicate the answer should be "yes" (assuming, at least, that the lay-legal concept of reasonable care held in a community does not shift over time into alignment with the legal concepts pronounced by judges, and therefore that the desired alignment is more plausibly achieved by judges adopting lay views than by laypeople adopting judicial views. Given empirical evidence that lay views of "reasonable expectations of privacy have scarcely shifted toward alignment with the judicial views reflected in prominent Fourth Amendment decisions of the U.S. Supreme Court," *see, e.g.*, Tonja Jacobi & Christopher Brett Jaeger, Katz*'s Imperfect*

As noted above, these normative arguments come with significant caveats. I acknowledge that whether they are persuasive hinges on one's view of the primary goals of tort law. I do not expect this argument will sway a theorist who believes that tort law's guiding star ought to be economic efficiency. If the Hand Formula conception of reasonable care leads to more efficient results than conceptions based on community understandings and norms, then the economic theorist will contend that the Hand Formula should trump those alternatives, no matter whether the Hand Formula aligns with lay understanding.

Regardless of whether one finds these normative arguments persuasive, my findings have notable implications for tort practitioners. In presenting negligence cases, trial lawyers make myriad choices about what evidence to present and how to present it. Tort lawyers know the Hand Formula.[305] While, as a strategic matter, lawyers do not always embrace it—defense lawyers, for instance, are often loath to hang their case on the fact that making a more dangerous product was cost-justified[306]—plaintiffs' lawyers will generally be keen to put on any evidence that a cheap, easily taken precaution could have prevented the injury, that the defendant's conduct created a high likelihood of an accident, or that any accident would be expected to be severe.

The findings of this Article, if replicated with other scenarios and more externally valid methods such as courtroom simulations, would provide some guidance about which type of evidence is most likely to persuade jurors, and thus which type of evidence lawyers may want to emphasize most. Similarly, to the extent jurors' evaluations of reasonableness may be driven by different criteria than judges' evaluations (which one might hypothesize stick closer to the Hand balancing that gets so much attention in law schools), my findings may have implications for tort litigators deciding whether to pursue a bench trial or a jury trial. For example, if your case fares better under Kantian criteria than under the

Circle: An Empirical Study of Reasonable Expectations of Privacy, 77 FLA. L. REV. 593, 614-16, 635 (2025), it seems quite unlikely that decisions of (typically state) courts would have substantial influence on lay understanding of reasonable care.).

On the other hand, there are serious questions about whether judges are (1) likely to successfully identify the lay-legal concept and (2) conform their reasoning to approximate that concept. Even assuming a judge could perform these challenging tasks, the judge's solo decision would still by definition lack the "local objectivity" that arises from consensus among people with varying normative perspectives, so the notion of "local objectivity" does not seem to cut one way or another on the bench-trial question. *See* Wells, *supra* note 298, at 2409. Further, concerns for the democratic values embodied in local jury decision-making are also less weighty if both parties to a dispute freely choose to have it decided by a judge rather than the jury. Ultimately, I remain agnostic as to this bench-trial question, which I believe merits deeper consideration.

305. *See* Kelley, *supra* note 34, at 732 & n.4.

306. Hetcher, *supra* note 114, at 649-50 ("Seasoned trial lawyers intuit that typical jurors would find a cost-benefit mode of reasoning morally unattractive.").

Hand Formula, a jury trial may be a better bet for you. Further, while the reality of modern practice is that most cases settle before trial, settlement negotiations are informed by lawyers' beliefs about likely outcomes at trial. Factoring in the primacy of *P* in juror decision-making might help lawyers better value negligence cases for settlement purposes—or provide a powerful argument for preferring one valuation to another.

One issue relevant to tort practitioners that could be explored in future work is how framing the same information in high- or low-*P* terms affects juror reasonableness judgments. For example, imagine a negligent-design case involving a product used by 100,000 people per year and an allegedly negligent design choice that increases the risk of a given user being injured by 0.1% over the next year. While defense counsel may prefer to speak in terms of this low-*P* framing, the plaintiff would likely benefit from a high-*P* framing—"100 more expected injuries per year." This framing—emphasizing the defendant's imposition of unaccepted costs, with a high degree of certainty and on a substantial number of others—smacks of a defendant treating others as means, not ends.[307]

C. Limitations

Throughout this Article, I have noted some of the methodological limitations of my experiments and potential hurdles to drawing conclusions about real-world tort decision-making. External validity and generalizability are always a concern with experimental research: can we expect that the results we observe in an experiment will manifest with different participants, in different situations, or within different settings? This Section discusses four limitations relating to the external validity and generalizability of my experiments.

First, my participants judged written vignettes presented online, not live courtroom cases with real parties and real stakes. A skeptical reader may question whether the factors that influence participants' reasonableness judgments in brief written vignettes generalize to real jurors judging real cases with real consequences. This is a legitimate concern; it is premature to jump from the results of my vignette studies to confident prognostications about the determinants of real-world litigation. But at the very least, an easily and relatively cheaply administered vignette-based study can point researchers toward hypotheses worth testing with more realistic (and more expensive) methods.[308] And it is

307. *Cf.* WARD FARNSWORTH & MARK F. GRADY, TORTS 441-42 (3d ed. 2019) (discussing reports from jurors in the famous "hot coffee" case, *Liebeck v. McDonald's*, No. CV-93-02419, 1995 WL 360309 (N.M. Dist. Aug. 18, 1994), suggesting that the jurors were troubled by McDonald's expert witness trivializing seven hundred complaints about "hot coffee" made over the course of a decade).

308. *See infra* Section III.D.

noteworthy that the research that has been done on juror studies suggests that results are typically similar across presentation formats; the patterns that emerge from studies with brief written vignettes tend to resemble those that emerge from more ecologically valid courtroom simulations.[309] A prominent meta-analysis of mock-juror studies "found little indication that . . . stimulus case realism (for example, paper-pencil versus audiotape versus videotape) . . . systematically influences research conclusions."[310]

Further, while it is true that mock-juror verdicts lack real consequences (no litigants are gaining or losing money), real-world jurors do not have any direct financial stake in the outcome of litigation either—the gains or losses are felt by the litigants, not the jurors. Jurors are typically paid a small amount for their service, but they get that money regardless of whether they find for the plaintiff or the defendant and whether they get the case "right" or "wrong." (Indeed, real-world jurors typically have no way to know for certain whether their decision was "right" or "wrong.") Thus, the incentives for mock jurors in research studies are not that different from jurors in real-world litigation. There is no obvious reason to expect that mock jurors in research settings would use entirely different criteria to judge reasonableness than jurors in real-world settings. While one might expect that mock jurors in studies may be less attentive than real-world jurors due to the lack of real stakes, if anything, a lack of attentiveness ought to make the criteria relevant to the mock jurors' decisions harder to identify. That is, if research participants are less attentive than real jurors, one would expect their judgments to be noisier, which suggests that the considerations that influence mock jurors' judgments might have even clearer influences on real-world jurors.

Second, my data came from mock jurors rather than mock juries.[311] In each experiment, participants responded as individuals rather than as part of groups. Thus, my participants did not engage in the interpersonal deliberative processes—and did not experience the group decision-making dynamics—that can affect real-world juries.[312] This is another important limitation that could be addressed in future research. That said, research studying group decision-making processes in mock juries has revealed that (1) most jurors tend to have verdicts

309. *See* Robert J. MacCoun, *Comparing Legal Factfinders: Real and Mock, Amateur and Professional*, 32 FLA. ST. U. L. REV. 511, 512-15 (2005).

310. *Id.* at 512 (arguing that "efforts to maximize realism . . . have more to do with research marketing than scientific validity").

311. *See* REBECCA K. HELM, HOW JURIES WORK 141 (2025) ("[S]ignificantly less research has examined jury verdicts compared to individual juror verdicts, [though] research highlights the importance of juror verdicts in predicting jury verdicts.").

312. *See id.* at 141-53 (reviewing research on "how individual juror impressions are translated into jury verdicts").

in mind before starting deliberations;[313] (2) most jurors tend to stick to that predeliberation verdict;[314] (3) larger factions tend to prevail in jury deliberations, with "the relationship between individual verdict preferences at the start of deliberation and the jury's verdict best described by a proportionality process in which the initial proportion of jurors favoring a verdict is directly related to the probability of the jury reaching that verdict";[315] and relatedly, (4) jurors who shift away from their predeliberation verdicts tend to shift toward the largest coalition. For instance, if a jury of six contains one juror who enters deliberations leaning "not liable" and five leaning "liable," the "not liable" juror is more likely to shift to "liable" than one of the "liable" jurors is to shift to "not liable."[316] While a variety of group processes could undoubtedly affect the results of particular mock juries, there is reason to expect that patterns in jurors' individual leanings before (or without) deliberation will generally tend to manifest in jury verdicts, too.

Third, representativeness of the participants in the experimental sample(s) is always a concern. Like all experimental data, my data come from samples—and samples can never be in all ways perfectly representative of the relevant population from which they are drawn. For example, my samples tended to be more female and more educated than the general population of Americans: of the 1,009 participants who identified as either male or female, 56.9% identified as female, and 59.2% of my participants reported having at least a college degree. It is, in theory, possible that people who identify as female or people with more education are more inclined to conceptualize tort law's reasonable-care standard in, say, economic terms than are other people, and thus the nonrepresentativeness of my sample could affect the generalizability of my results to other samples.[317] Concerns along these lines, however, are ameliorated at least in part by

313. *See, e.g.*, Reid Hastie, David A. Schkade & John W. Payne, *A Study of Juror and Jury Judgments in Civil Cases: Deciding Liability for Punitive Damages*, 22 LAW & HUM. BEHAV. 287, 295 (1998) (finding that, overall, 91% of participants reported a predeliberation verdict of liability or nonliability, while only 9% reported being undecided).

314. *Id.* at 296 (finding that only 23% of participants with definite predeliberation verdicts switched verdicts through the deliberation process).

315. *Id.* at 305; *see also* HELM, *supra* note 311, at 141 ("Very broadly speaking, research suggests that the probability of a jury reaching a particular verdict is roughly proportional to the initial number of jurors favouring that verdict.").

316. Hastie et al., *supra* note 313, at 297, 305.

317. There is also the possibility that participants recruited through Prolific (as my participants were) might differ from the general population of jury-eligible Americans, or from the subset of jury-eligible Americans most likely to serve on juries. *See* Shari Seidman Diamond & Mary R. Rose, *Real Juries*, 1 ANN. REV. L. & SOC. SCI. 255, 257 (2005) (noting "the persistent reality . . . that the prospective jurors assembled in courthouses and the empanelled juries that

the fact that my data do not indicate that there are any significant, consistent differences in negligence liability judgments based on participants' age, gender identity, racial or ethnic identity, or education level (nor am I aware of any theory that would predict such differences). The Appendix provides, for each experiment, a breakdown of the gender, race, ethnicity, age, and education-level data collected from participants, as well as a summary of exploratory analyses conducted using this demographic information.[318]

Fourth, and finally, my studies asked participants to judge reasonableness in particular fact patterns; therefore, my results may not generalize to the full universe of tort cases. Most obviously, all of my vignettes involved cases where participants had full, unambiguous information about the values of *B*, *P*, and *L*—a situation that is rare in real-world litigation.[319] Consequently, my studies might overestimate the influences, both economic and Kantian, of *B*, *P*, and *L* relative to typical real-world cases in which information about these inputs is less clear to jurors or is contested by parties (and therefore could factor less prominently in juror decisions). One natural avenue for future research—discussed at more length in Section III.D, *infra*—is to explore whether lay decision-makers explicitly or implicitly make estimates of *B*, *P*, and *L* absent unambiguous and uncontested values and how such estimates affect negligence judgments. While I acknowledge that the design of my studies could lead them to overestimate the role of Kantian or economic reasoning relative to real-world cases, it is very unlikely that my studies underestimate the role of Kantian or economic reasoning relative to these real-world scenarios. For example, if the Hand Formula were ever going to perfectly capture lay negligence judgments, it seems it would be in the contexts that I tested. It seems implausible that participants would be more likely to strictly apply the Hand Formula to cases without clear and undisputed *B*, *P*, and *L* values than they would to cases with clear and undisputed *B*, *P*, and *L* values.[320]

---

decide cases are systematically different from the communities from which they were drawn"). That said, I am unaware of any reason to suspect systematic differences across these groups with respect to what criteria they use to judge reasonableness.

318. Future research might empirically examine whether and how negligence judgments vary across demographic groups. For a general discussion, see Mikaela Spruill & Neil A. Lewis, Jr., *How Do People Come to Judge What Is 'Reasonable'? Effects of Legal and Sociological Systems on Human Psychology*, 18 PERSPS. ON PSYCH. SCI. 378, 378 (2023). This was not a focus in the experiments reported in this Article, and my sample sizes were predicated on analyses of the full group of participants (rather than on analyses of subsamples based on demographic characteristics). Post hoc exploration of my data did not provide any evidence of significant differences in negligence judgments based on participants' age, gender identity, racial identity, or education level. *See infra* Appendix.

319. *See supra* Part II.

320. *See supra* notes 164-168 and accompanying text.

It is also possible that the particular case vignettes used in my studies led participants to consider *B*, *P*, and *L* information in ways they would not have considered these datapoints in other cases. While I deliberately varied my fact pattern from experiment to experiment to cover different types of tort cases (e.g., property damage versus personal injury; individual defendant versus corporate defendant), there are countless other dimensions along which tort cases vary. For example, was the defendant pursuing a private or public benefit when the defendant injured the plaintiff?[321] Was the defendant subjectively aware of the risk of injury to the plaintiff?[322] More research is needed to assess the breadth of the influence of economic and Kantian concerns across different types of tort cases.

D. Future Directions for Research

This paper contributes to the growing field of experimental jurisprudence, which endeavors to "address jurisprudential questions with empirical data."[323] The jurisprudential debate surrounding reasonableness is especially well suited for empirical work. This Section highlights some questions raised by my findings that can be addressed through future studies.

As discussed in Section III.C, vignette-based legal research naturally invites follow-up work to address generalizability and external-validity limitations. For example, my vignettes (or similar vignettes) could be presented in forms that better approximate real-world litigation. There is a spectrum of available approaches, ranging from simulated trial transcripts laying out each side's case to audio or video recordings to full jury simulations. Such work would be valuable in building a bridge from the lab to litigation.

Following up on my findings, the result that *P* has an outsized influence among the Hand Formula inputs invites research about the nature of its influence. For example, is the relation between *P* and negligence judgments linear, meaning that increases in *P* lead to consistent, corresponding increases in perceived negligence? Or are there "jumps" in perceived negligence when *P* crosses a certain threshold that feels particularly "significant" (e.g., 10%, or 50%)? Analogous questions could be asked about the effect of the product *PL*.[324] Another useful line of inquiry: when the value of *P* is not directly presented to research participants or jurors, do they implicitly or explicitly infer a value of *P*? If they

---

321. *See* Stein, *supra* note 289, at 541.

322. *See* Margoni & Brown, *supra* note 125, at 7.

323. Kevin Tobia, *Experimental Jurisprudence*, 89 U. CHI. L. REV. 735, 736 (2022).

324. Analogous questions could also be asked about *B* and *L*, for that matter.

do infer a *P* value, does it have the same outsized influence on their judgments as when it is directly presented?

The influence of *B* in my studies also merits further investigation. One question of great interest is how much of *B*'s influence is attributable to its role in Hand-style cost-benefit analysis versus a potential *B*-focused heuristic.[325] Future research could probe this question using my design from Experiments Four and Five by inverting the relationship between the values experimentally assigned to *B*, *P*, and *L*. That is, researchers could assign values of *B* and *PL* such that the low *B* value is greater than the low *PL* value and the high *B* value is less than the high *PL* value. With values assigned in this manner, *PL* (rather than *B*) would become the stand-in for cost justification of the Hand Formula: the Hand Formula would indicate the defendant behaved unreasonably whenever *PL* was high and reasonably whenever *PL* was low. With this approach, any effect of manipulating *B* would be independent of its role in the Hand Formula, reflecting an irrational, *B*-focused heuristic (as opposed to rational Hand balancing).

With respect to *L*, I deliberately avoided any scenarios that created a risk of, or resulted in, wrongful death. I designed my studies with both *B* and *L* quantified in dollars, so that the Hand Formula would generate unambiguous answers about whether the defendant acted reasonably. Given the reticence of people to place dollar values on human lives, I avoided scenarios involving death out of concern that people would reject the quantification of *L*. But future research could manipulate *L* such that some conditions involve the risk of death and others do not.[326] Doing so may reveal that in such cases *L*, like *P*, has a Kantian effect above and beyond its role in the Hand Formula.

Finally, my broadest call for further research relates to the variance in lay negligence ratings that was not explained by *B*, *P*, and *L* information in my studies. My findings provide evidence that both economic and Kantian concerns factor into lay negligence judgments, but they leave much of the variance in those judgments unexplained. Future work can and should investigate other legal and behavioral theories that factor into lay negligence judgments.

---

325. *See supra* Section III.A.3.

326. Indeed, *L* could be still more catastrophic than one death. When scientists working on the Large Hadron Collider prepared for cutting-edge experiments, there was a very small risk the experiments would create a black hole that would swallow the Earth and all of humanity with it—an infinite *L* for a very-small-*P* scenario. *See* Eric E. Johnson, *The Black Hole Case: The Injunction Against the End of the World*, 76 TENN. L. REV. 819, 877-78 (2008).

CONCLUSION

A defendant's liability for tort negligence often turns on whether jurors decide the defendant acted reasonably or unreasonably. But how do jurors make this decision? Which criteria do they use? To the extent it has addressed this question, legal scholarship has often treated it as an all-or-nothing issue—the reasonableness standard is defined only by the Hand Formula, or only by whether the defendant has imposed a "significant" risk on others without consent. But examining the reasonableness debate through the lens of cognitive psychology suggests it may be more useful to think of reasonableness judgments as functions of multiple inputs. From this perspective, the issue is not whether lay decision-makers think in purely economic or purely Kantian terms. The answer is surely neither. A psychological perspective sets up a more tractable question: to what extent do economic and Kantian considerations—along with other types of considerations—affect what laypeople judge to be reasonable?

Working from this perspective, this Article makes two key contributions. First, it translates prominent reasonableness theories into tractable predictions about what types of information will influence reasonableness judgments in the context of garden-variety tort disputes. Second, my experiments tested those predictions and found evidence that both economic and Kantian inputs factor into the function of reasonableness. Lay decision-makers judged defendants more negligent when *B* was less than *PL* than when *B* was greater than *PL*, as one would expect if they engaged (consciously or unconsciously) in some balancing of costs and benefits. But lay decision-makers were also influenced by the significance of these risks separate and apart from their role in the Hand Formula. My studies demonstrate that *PL*, and especially *P*, loom larger than the Hand Formula suggests, reflecting a distinctively Kantian influence on reasonableness.

APPENDIX

This Appendix provides detailed information about the designs of my experiments and demographic information about the participants sampled for each. It also provides information about supplemental post hoc analyses examining the influences of participants' demographic characteristics on negligence ratings. In short, these post hoc analyses provide little evidence that there are consistent differences across demographic groups in evaluating negligence. It is important to note, however, that my studies were not designed to test for such differences — future work, with larger samples or different sampling methods, might further investigate whether there are differences between demographic groups.

A. Experiment One

APPENDIX TABLE 1. SUMMARY OF DESIGN FOR EXPERIMENT ONE (PROCEEDS ON FOLLOWING TWO PAGES)

Experimental Manipulations			Negligence Rating Predictions Based on Theory		
B	*P*	*L*	*Theory:* Economic *Criteria: B < PL*	*Theory:* Strong Kantian (Broad) *Criteria: PL*	*Theory:* Strong Kantian (Narrow) *Criterion: P*
Low ($6)	Low (0.2%)	Low ($4K)	*Prediction:* High negligence rating *Reason:* *B* ($6) < *PL* ($8)	*Prediction:* Low negligence rating *Reason:* Low *PL* ($8)	*Prediction:* Low negligence rating *Reason:* Low *P* (0.2%)
Low ($6)	Low (0.2%)	High ($400K)	*Prediction:* High negligence rating *Reason:* *B* ($6) < *PL* ($800)	*Prediction:* Intermediate negligence rating *Reason:* Intermediate *PL* ($800)	*Prediction:* Low negligence rating *Reason:* Low *P* (0.2%)
Low ($6)	High (20%)	Low ($4K)	*Prediction:* High negligence rating *Reason:* *B* ($6) < *PL* ($800)	*Prediction:* Intermediate negligence rating *Reason:* Intermediate *PL* ($800)	*Prediction:* High negligence rating *Reason:* High *P* (20%)
Low ($6)	High (20%)	High ($400K)	*Prediction:* High negligence rating *Reason:* *B* ($6) < *PL* ($80K)	*Prediction:* High negligence rating *Reason:* High *PL* ($80K)	*Prediction:* High negligence rating *Reason:* High *P* (20%)

Experimental Manipulations			Negligence Rating Predictions Based on Theory		
B	*P*	*L*	*Theory:* Economic *Criteria: B < PL*	*Theory:* Strong Kantian (Broad) *Criteria: PL*	*Theory:* Strong Kantian (Narrow) *Criterion: P*
High ($100K)	Low (0.2%)	Low ($4K)	*Prediction:* Low negligence rating *Reason:* *B* ($100K) > *PL* ($8)	*Prediction:* Low negligence rating *Reason:* Low *PL* ($8)	*Prediction:* Low negligence rating *Reason:* Low *P* (0.2%)
High ($100K)	Low (0.2%)	High ($400K)	*Prediction:* Low negligence rating *Reason:* *B* ($100K) > *PL* ($800)	*Prediction:* Intermediate negligence rating *Reason:* Intermediate *PL* ($800)	*Prediction:* Low negligence rating *Reason:* Low *P* (0.2%)
High ($100K)	High (20%)	Low ($4K)	*Prediction:* Low negligence rating *Reason:* *B* ($100K) > *PL* ($800)	*Prediction:* Intermediate negligence rating *Reason:* Intermediate *PL* ($800)	*Prediction:* High negligence rating *Reason:* High *P* (20%)
High ($100K)	High (20%)	High ($400K)	*Prediction:* Low negligence rating *Reason:* *B* ($100K) > *PL* ($80K)	*Prediction:* High negligence rating *Reason:* High *PL* ($80K)	*Prediction:* High negligence rating *Reason:* High *P* (20%)

APPENDIX TABLE 2. DESCRIPTIVE STATISTICS OF PARTICIPANTS IN EXPERIMENT ONE

Category	N	Percentage
Gender		
Female	137	64.0
Male	68	31.8
Nonbinary/Nonconforming	7	3.3
Prefer Not to Respond	1	0.5
Data Missing (Participant did not choose any of these options.)	1	0.5
Racial/Ethnic Identity		
American Indian or Alaska Native	1	0.5
Asian	12	5.6
Black or African American	20	9.3
Hispanic, Latino, or of Spanish Origin	9	4.2
Middle Eastern or North African	0	0.0
Native Hawaiian or Pacific Islander	0	0.0
White or Caucasian	153	71.5
Multiethnic/Multiple Categories from Above	19	8.9
Prefer Not to Disclose	0	0.0
Education level		
Some High School	2	0.9
High School Degree	31	14.5
Some College	43	20.1
College Degree	103	48.1
Master's Degree	16	7.5
Professional Degree (J.D., M.D., etc.)	9	4.2
Graduate Degree (Ph.D., etc.)	10	4.7
Age		
18-30	47	22.0
31-40	62	29.0
41-50	39	18.2
51-60	33	15.4
61-70	28	13.1
71+	4	1.9
Data Missing (Participant did not choose any of these options.)	1	0.5

Participants' negligence ratings did not differ significantly by gender or by race. Further, there was no correlation between age and negligence rating or between parent education level and negligence rating. There was, however, a significant correlation between the participants' own education level and negligence, with higher education levels correlating with lower negligence ratings ($r = -.195$, $p = .004$).

Rerunning my primary analysis as an ANCOVA incorporating as covariates (1) age, (2) a numerically coded variable for male or female gender identity (male = 1, female = 0), (3) a numerically coded variable for education level (from some high school = 1 to graduate degree = 7), and (4) a numerically coded variable for parent education level (from some high school = 1 to graduate degree = 7) yielded the same fundamental pattern of results as my primary analysis: *B* information and *P* information significantly influenced negligence ratings (both p's < .001), and *L* information did not, with no interactions. One covariate—the participant's education level—was a significant predictor in this ANCOVA model ($p = .005$).

B. Experiment Two

APPENDIX TABLE 3. DESCRIPTIVE STATISTICS OF PARTICIPANTS IN EXPERIMENT TWO

Category	N	Percentage
Gender		
Female	113	56.8
Male	80	40.2
Nonbinary/Nonconforming	3	1.5
Prefer not to respond	0	0
Data Missing (Participant did not choose any of these options.)	3	1.5
Racial/Ethnic Identity		
American Indian or Alaska Native	0	0
Asian	12	6.0
Black or African American	14	7.0
Hispanic, Latino, or of Spanish Origin	6	3.0
Middle Eastern or North African	0	0
Native Hawaiian or Pacific Islander	0	0
White or Caucasian	152	76.4
Multiethnic/Multiple Categories from Above	13	6.5

Prefer Not to Disclose	2	1.0
Education Level		
Some High School	2	1.0
High School Degree	24	12.1
Some College	52	26.1
College Degree	101	50.8
Master's Degree	13	6.5
Professional Degree (J.D., M.D., etc.)	4	2.0
Graduate Degree (Ph.D., etc.)	3	1.5
Age		
18-30	61	30.7
31-40	49	24.6
41-50	49	24.6
51-60	19	9.5
61-70	20	10.1
71+	1	0.5

Participants' negligence ratings did not differ significantly by gender or by race. Further, there were no significant correlations between age and negligence rating, between education level and negligence rating, or between parent education level and negligence rating.

Rerunning my primary analysis as an ANCOVA incorporating as covariates (1) age, (2) a numerically coded variable for male or female gender identity (male = 1, female = 0), (3) a numerically coded variable for education level (from some high school = 1 to graduate degree = 7), and (4) a numerically coded variable for parent education level (from some high school = 1 to graduate degree = 7) yielded results broadly similar but not identical to my primary analysis: *B* information and *L* information again significantly influenced negligence ratings (both p's < .001), though in the ANCOVA model the influence of *P* information fell just short of statistical significance (p = .088). Further, in the ANCOVA model, there was a significant *B*P* interaction (p = .037); no other interactions were significant.

C. Experiment Three

APPENDIX TABLE 4. DESCRIPTIVE STATISTICS OF PARTICIPANTS IN EXPERIMENT THREE

Category	N	Percentage
Gender		
Female	108	51.7
Male	93	44.5
Nonbinary/Nonconforming	2	1.0
Prefer Not to Respond	1	0.5
Data Missing (Participant did not choose any of these options.)	5	2.4
Racial/Ethnic Identity		
American Indian or Alaska Native	1	0.5
Asian	18	8.6
Black or African American	14	6.7
Hispanic, Latino, or of Spanish Origin	4	1.9
Middle Eastern or North African	0	0
Native Hawaiian or Pacific Islander	0	0
White or Caucasian	153	73.2
Multiethnic/Multiple Categories from Above	17	8.1
Other	1	0.5
Prefer Not to Disclose	0	0
Data Missing (Participant did not choose any of these options.)	1	0.5
Education Level		
Some High School	1	0.5
High School Degree	29	13.9
Some College	62	29.7
College Degree	95	45.5
Master's Degree	19	9.1
Professional Degree (J.D., M.D., etc.)	2	1.0
Graduate Degree (Ph.D., etc.)	1	0.5
Age		
18-30	41	19.6
31-40	68	32.5
41-50	40	19.1
51-60	33	15.8
61-70	17	8.1
71+	10	4.8

Participants' negligence ratings did not differ significantly by gender or by race. Further, there were no significant correlations between age and negligence rating, between education level and negligence rating, or between parent education level and negligence rating.

Rerunning my primary analysis as an ANCOVA incorporating as covariates (1) age, (2) a numerically coded variable for male or female gender identity (male = 1, female = 0), (3) a numerically coded variable for education level (from some high school = 1 to graduate degree = 7), and (4) a numerically coded variable for parent education level (from some high school = 1 to graduate degree = 7) yielded the same fundamental pattern of results as my primary analysis: *P* information significantly influenced negligence ratings (p = .002), while *B* information and *L* information did not (with no interactions).

D. Experiment Four

APPENDIX TABLE 5. SUMMARY OF DESIGN FOR EXPERIMENT FOUR

Experimental Manipulations		**Negligence Rating Predictions Based on Theory**	
B	*PL*	*Theory:* Economic *Criteria: B* < *PL*	*Theory:* Strong Kantian *Criteria: P* (Narrow) or *PL* (Broad)
Low ($90)	Low *P* (0.6%) *L* = $22.5K	*Prediction:* High negligence rating *Reason: B* ($90) < *PL* ($135)	*Prediction:* Low negligence rating *Reason:* Low *P* (0.6%), *PL* ($135)
Low ($90)	High *P* (60%) *L* = $10K	*Prediction*: High negligence rating *Reason: B* ($90) < *PL* ($6K)	*Prediction:* High negligence rating *Reason:* High *P* (60%), *PL* ($6K)
High ($9K)	Low *P* (0.6%) *L* = $22.5K	*Prediction*: Low negligence rating *Reason: B* ($9K) > *PL* ($135)	*Prediction:* Low negligence rating *Reason:* Low *P* (0.6%), *PL* ($135)
High ($9K)	High *P* (60%) *L* = $10K	*Prediction:* Low negligence rating *Reason: B* ($9K) > *PL* ($6K)	*Prediction:* High negligence rating *Reason:* High *P* (60%), *PL* ($6K)

APPENDIX TABLE 6. DESCRIPTIVE STATISTICS OF PARTICIPANTS IN EXPERIMENT FOUR

Category	N	Percentage
Gender		
Female	119	55.3
Male	90	41.9
Nonbinary/Nonconforming	3	1.4
Prefer Not to Respond	0	0
Data Missing (Participant did not choose any of these options.)	3	1.4
Racial/Ethnic Identity		
American Indian or Alaska Native	0	0
Asian	17	7.9
Black or African American	28	13.0
Hispanic, Latino, or of Spanish Origin	5	2.3
Middle Eastern or North African	2	0.9
Native Hawaiian or Pacific Islander	0	0
White or Caucasian	144	67.0
Multiethnic/Multiple Categories from Above	18	8.4
Other	1	0.5
Prefer Not to Disclose	0	0
Education level		
Some High School	2	0.9
High School Degree	26	12.1
Some College	59	27.4
College Degree	96	44.7
Master's Degree	23	10.7
Professional Degree (J.D., M.D., etc.)	7	3.3
Graduate Degree (Ph.D., etc.)	2	0.9
Age		
18-30	67	31.2
31-40	70	32.6
41-50	42	19.5
51-60	26	12.1
61-70	6	2.8
71+	4	1.9

Participants' negligence ratings did not differ significantly by gender or by race. Further, there were no significant correlations between age and negligence rating, between education level and negligence rating, or between parent education level and negligence rating.

Rerunning my primary analysis as an ANCOVA incorporating as covariates (1) age, (2) a numerically coded variable for male or female gender identity (male = 1, female = 0), (3) a numerically coded variable for education level (from some high school = 1 to graduate degree = 7), and (4) a numerically coded variable for parent education level (from some high school = 1 to graduate degree = 7) yielded the same fundamental pattern of results as my primary analysis: *P* information ($p < .001$) and *B* information ($p = .024$) both significantly influenced negligence ratings (with no interaction).

E. Experiment Five

APPENDIX TABLE 7. DESCRIPTIVE STATISTICS OF PARTICIPANTS IN EXPERIMENT FIVE

Category	N	Percentage
Gender		
Female	97	46.4
Male	104	49.8
Nonbinary/Nonconforming	6	2.9
Prefer Not to Respond	0	0
Data Missing (Participant did not choose any of these options.)	2	1.0
Racial/Ethnic Identity		
American Indian or Alaska Native	0	0
Asian	18	8.6
Black or African American	19	9.1
Hispanic, Latino, or of Spanish Origin	4	1.9
Middle Eastern or North African	2	1.0
Native Hawaiian or Pacific Islander	0	0
White or Caucasian	138	66.0
Multiethnic/Multiple Categories from Above	26	12.4
Prefer Not to Disclose	2	1.0
Education Level		
Some High School	3	1.4
High School Degree	40	19.1

Some College	51	24.4
College Degree	93	44.5
Master's Degree	16	7.7
Professional Degree (J.D., M.D., etc.)	2	1.0
Graduate Degree (PhD, etc.)	4	1.9
Age		
18-30	63	30.1
31-40	76	36.4
41-50	27	12.9
51-60	25	12.0
61-70	13	6.2
71+	5	2.4

Participants' negligence ratings did not differ significantly by gender or by race. Further, there were no significant correlations between age and negligence rating, between education level and negligence rating, or between parent education level and negligence rating.

Rerunning my primary analysis as an ANCOVA incorporating as covariates (1) age, (2) a numerically coded variable for male or female gender identity (male = 1, female = 0), (3) a numerically coded variable for education level (from some high school = 1 to graduate degree = 7), and (4) a numerically coded variable for parent education level (from some high school = 1 to graduate degree = 7) yielded the same fundamental pattern of results as my primary analysis: *P* information ($p < .001$) and *B* information ($p < .001$) both significantly influenced negligence ratings, with a significant $B*P$ interaction ($p = .001$). One covariate—parent education level—was a significant predictor in this ANCOVA model ($p = .009$).