

‘I Didn’t Know It Was You’: The Impersonal Grounds of Relational Normativity*

Jed Lewinsohn (jed@pitt.edu)

[Forthcoming in *Noûs*; please cite final version]

Abstract: A notable feature of our moral and legal practices is the recognition of privileges, powers, and entitlements belonging to a select group of individuals in virtue of their status as victims of wrongful conduct. A philosophical literature on relational normativity purports to account for this status in terms of such notions as interests, rights, and attitudes of disregard. This paper argues that such individualistic notions cannot account for prevailing and intuitive ways of demarcating the class of victims. The paper is focused on the wrongful infliction of harm, and centers on the mediating role played by impersonal “danger-making properties” in the determination of the class of victims. The paper begins with an analysis of one of the most well-known discussions of negligently-inflicted harm — from the most famous case of the American common law tradition — and the analysis is then extended to the morality of harm-doing more broadly, negligent and intentional alike. The paper’s chief targets are interest theories of rights — including contractualist theories of moral claim-rights of the kind defended by R. Jay Wallace — and neo-Strawsonian Quality of Will theories of “moral injury”.

With respect to many kinds of bad behavior, our moral and legal practices endow upon a select group of individuals the status of the *properly aggrieved*. When it comes to matters such as forgiveness, apology, and duties of repair, members of this set - the set of victims – are commonly thought to stand in a privileged position relative to others. Such privilege may even extend to matters of the heart: certain moral emotions, such as seething resentment, are sometimes thought to be appropriate feelings for the properly aggrieved, but not for others. It is my contention that this privileging of a select group of victims cannot be accounted for using the highly individualistic notions that have dominated philosophical discussions of relational normativity. In particular, I will argue that membership in the class of the properly aggrieved is explained neither by a wrongful act’s impingement or infringement upon the *interests* or *rights* of class members, nor by the act’s manifestation of *disregard* or *disrespect* for them. The

* For generous feedback I am grateful to Sam Berstler, James Brandt, Ben Eidelson, John Goldberg, Eli Hirsch, Felix Koch, Crescente Molina, Ezequiel Monti, Emilio Mora, Richard Moran, Ketan Ramakrishnan, Filippa Ronquist, Jacob Rosen, Nick Sage, Erica Shumener, Adam Slavny, Henry Smith, and an anonymous reviewer, as well as to audiences at Berkeley, Harvard, MIT, and the Philosophy & Private Law Discussion Group.

alternative I favor is that such a determination is at bottom a product of social and institutional accountability practices, and should be evaluated accordingly.¹

The focus of my discussion will be the wrongful infliction of harm, a domain in which the identification of victims by appeal to their interests or their rights may seem especially promising. After all, when an act's wrongfulness is explained by its adverse effects on the interests of certain individuals, it is natural to think that the same explanation straightforwardly identifies the individuals who are properly aggrieved on account of the act. Although my primary interest is in relational morality, I will begin by considering what is surely the greatest hit of the American common law tradition and one of the most well-known discussions of negligently-inflicted harm. I will then extend the analysis (in each section) from law to morality and from negligence to intentional harm-doing.

The common law case I will consider concerns the relation between two basic limitations on an actor's liability for negligently-inflicted harm. The first limitation states that an individual is liable for losses resulting from their negligent conduct only if they violated a duty of care owed to the injured party. It is not enough, in other words, for the injured party to show that the negligent actor violated a duty of care owed to somebody else, or to nobody in particular. The second limitation provides that an injured party is entitled to recover from a negligent actor only

¹ Terminological note: as I am using the term, "proper aggrievement" refers to a *status* that results from one's relation to another person's wrongful or culpable conduct and that *presumptively* endows one with certain powers, privileges, and entitlements. And while some use the expression "X was wronged" to refer to proper aggrievement (in my sense), I use it to mean "someone violated an obligation owed to X."

if the injured party belongs to the class of individuals who were foreseeably endangered by the negligent act. It is the central assumption of Benjamin Cardozo's celebrated opinion in *Palsgraf v. Long Island Railroad Company* that the second limitation follows from the first — that is, that a negligent act violates a duty of care owed to a particular individual only if the individual belongs to the class of individuals foreseeably endangered by the negligent act.²

In *Palsgraf*, a guard working for the defendant's railroad, in an effort to help a passenger board an already moving train, pushed the passenger, with the foreseeable risk that the passenger would be harmed. It transpired that, as a result of the push, a package containing fireworks was dislodged from the passenger's hands, causing an explosion that injured the plaintiff, Helen Palsgraf, who was standing on the platform far away from the train. The guard could not have foreseen that the package would contain fireworks or that the plaintiff would be harmed as a result of the push. Cardozo, speaking for the majority, ruled in favor of the defendant. Cardozo reasoned that because Helen Palsgraf did not belong to the class of individuals foreseeably endangered by the guard's conduct, such conduct did not amount to a violation of a duty owed to her, and that absent such a violation, the defendant was not liable for her losses. "What the plaintiff must show is a 'wrong' to herself, i.e., the violation of her own right, and not merely a wrong to someone else, nor conduct 'wrongful' because unsocial, but not a 'wrong' to anyone."³ Leading philosophers of tort law have seized upon Cardozo's words in an effort to show that tort law's concern with the vindication of rights serves to explain not only its procedural dimensions,

² *Palsgraf v. Long Island Railroad*, 248 N.Y. 339 (1928). As Schwartz (1999, p. 312) notes, "it is *Palsgraf* that has come to symbolize the entire experience of being a law student."

³ *Id.*, pp. 343–44.

but also its substantive liability rules.⁴ However, a closer inspection of Cardozo’s reasoning reveals him to be relying on two assumptions that cannot be jointly held. First, by *grounding* his liability judgment in a determination that the defendant did not violate the plaintiff’s rights (i.e., did not violate an obligation owed to the plaintiff), Cardozo presupposes a conception of *violating X’s right* that is distinct from *being liable for X’s losses*.⁵ Second, Cardozo assumes (as already noted) that an individual’s rights are not violated by a negligent act when the individual does not belong to the class of individuals foreseeably endangered by it. However, once we properly characterize the latter class (i.e., the class of individuals foreseeably endangered by a negligent act), it will emerge that the only conception of rights that vindicates the second assumption is one on which the notions of *rights-violations* and *liabilities* coincide. And by extending the analysis to the moral case, we will come to see that what is true for legally-enforceable liabilities is true for proper aggrievement more generally.

I Three Conceptions of the Duty of Care

As we have just seen, the basis of the *Palsgraf* decision turned on the relational character of the duty of care — according to Cardozo, the defendant railroad was not responsible for the plaintiff’s losses on the ground that it did not violate a duty of care owed to the plaintiff. In this

⁴ See, e.g., Goldberg and Zipursky (2021, pp. 198–205); Zipursky (1998); Weinrib (2003, pp. 159–167); Ripstein (2016, pp. 87–92); Coleman (2008, p. 1159); Perry (2014, n. 11).

⁵ Certain legal scholars, unable to even conceive of a distinction between rights and liability, were quick to charge Cardozo with circular reasoning: “This is merely a definition in a circle, and has all the truth and all the sterility of every tautologous proposition” (Cowan, 1938); “That is merely a dog chasing its own tail” (Prosser, 1953).

section, I articulate three relational conceptions of the duty of care, each of which identifies a required course of conduct as well as a class of individuals (“right-holders”) to whom the duty is owed.⁶ One of these conceptions underpins Cardozo’s position that Helen Palsgraf was not wronged by the guard’s act, and can be attributed to tort law more broadly. The second conception is a variant of Cardozo’s — on plausible factual assumptions, it would yield the same outcome in *Palsgraf* — and serves to bring into relief several distinctive features of tort law’s characterization of the class of individuals foreseeably endangered by a negligent act. According to the final conception, a negligent actor wrongs *every* other person, regardless of whether they belong to the class of the foreseeably endangered — a view that is resonant with (one reading of) the dissenting opinion in the case, written by Judge Andrews. Teeing up these three relational conceptions will allow us to take up the question of whether the selection of right-holders belonging to Cardozo’s conception can claim the support of a coherent conception of rights.

Of course, there is an important sense in which all three conceptions can claim the support of a leading theory of rights, the “Will Theory” of rights, widely associated (nowadays) with H.L.A. Hart. According to this account, the notion of a legal right correlative to a certain legal duty can be analyzed in terms of a bundle of legal powers related to that duty — most notably, the power to enforce the duty by suing for damages or injunctive relief in the event of breach (Hart, 1982, pp. 183–84). While this conception of legal rights is consistent with Cardozo’s conception of the duty of care, it is not consistent with his argument for that

⁶ I will follow Cardozo in using “X owes a duty to Y” and “Y has a (claim-)right against X” interchangeably. None of my arguments turn on this choice, as the theories of rights to be considered may be transposed into theories of directed duties.

conception.⁷ For in arguing that Helen Palsgraf lacks the power to recover damages on the ground that her rights were not violated, Cardozo presupposes a notion of a legal right that is independent of such a power. More generally, insofar as the notion of a right is analyzed in terms of the powers, privileges and entitlements constitutive of what I have called proper aggrievement (viz., those pertaining to apology, repair, forgiveness, and resentment), the fact that somebody's right has been violated cannot serve to explain why they belong to the class of the properly aggrieved.⁸

Before stating the three conceptions of the duty of care, I must first make explicit several assumptions that will guide my formulations throughout this discussion: First, since it does not affect the analysis, I will not qualify the principles to reflect the important fact that one's prior dealings, special relationships, and social roles can have an impact on one's duties of care. Second, I will assume that the harm in question is physical damage to person or property. Third, in stating the different conceptions I follow tradition (as well as Cardozo and Andrews) in assuming that the breach of a duty of care does not entail harm — that is, I shall assume that

⁷ Hart would have concurred (1982, n. 74). It should be emphasized that in criticizing Cardozo's argument I am not criticizing his liability judgment.

⁸ The same point holds, *mutatis mutandis*, if one analyzes a claim-right in terms of a more expansive bundle of powers and privileges, e.g., those constituting a relation of authority and accountability that has a prospective dimension (e.g., a power to demand compliance) as well as a retrospective one. See, e.g., Darwall (2006, pp. 18–19); see also Ripstein (2016).

‘breach of duty’ and ‘harm’ constitute distinct ‘elements’ of a negligence claim.⁹ Although I will assume this, I do not rely on this assumption.

I will begin by stating Cardozo’s conception of the duty of care; for all the discussion that *Palsgraf* has received, I am not aware of any attempts to state Cardozo’s rule with any exactness. For reasons that will soon become clear, I label the rule “Danger Zone”:

Danger Zone: If an agent believes, or should believe, that by performing a certain act ϕ they would create an unreasonable risk of harming someone or other who instantiates a certain “danger-making” property F (e.g., the property of occupying a certain spatiotemporal location or of drinking from a certain water source), then the agent owes a duty of care to all and only those who will even unforeseeably instantiate F not to ϕ .¹⁰

I will first clarify the rule in a number of respects, and then state the basis of the attribution to Cardozo. First, the relevant conception of risk is epistemic, not metaphysical. To better capture this, some may prefer to adopt the terminology of credences, or assigned probabilities: on this way of speaking, the antecedent concerns the credence that the agent either did assign or should have assigned, given the evidence available to her, to the proposition that her act would harm

⁹ Cardozo: “We are told that one who drives at reckless speed through a crowded city street is guilty of a negligent act and, therefore, of a wrongful one irrespective of the consequences. [Such an act is indeed] *wrongful...in relation to other travelers...*[to the extent that] the eye of vigilance perceives the risk of damage [to them].”

¹⁰ The antecedent of *Danger Zone* should be construed as entailing an unreasonable risk of harming someone or other (full stop), i.e., as taking into account the probability that the danger-making property will be instantiated.

someone or other instantiating a certain property. Second, the relevant notion of risk imports substantial uncertainty. Once the assigned probabilities rise to the level of “substantial certainty,” we depart from the field of negligence and enter the domain of intentional wrongdoing.¹¹

Likewise, *Danger Zone* does not apply if the agent aimed to cause the harm. Such an aim would again move us away from negligence to the intentional torts, where the liability rules do not turn in the same way on the reasonableness of the harmful act. For example if, to save my life, I deliberately destroy \$10 worth of your property, I must pay you back, notwithstanding the reasonableness of my act. By contrast, if, to save my life, I expose you to a 10% chance of getting killed, my conduct might be deemed reasonable, in which case I would not be liable in the event that the risk materializes. Third, while an analysis of what makes an action “unreasonably” (alternatively, “unduly”) risky lies beyond the scope of this paper, any adequate theory must in some way take into account the following features: the likelihood of harm occurring as a result of the act, the magnitude of that harm, and the costs (to the agent or others) of acting differently or of taking risk-mitigating precautions.

Fourth, whether a property qualifies as danger-making, relative to a certain action, is not conclusively settled by determining whether the agent believes or should believe that her act carries an unreasonable risk of harming individuals who instantiate the property. In addition, the “danger-making” qualifier is introduced to incorporate the following three constraints that limit

¹¹ As Seavey wrote long ago, in an article cited in Cardozo’s *Palsgraf* opinion, “Risk, then, would seem to include the advertence of someone to the possibility that an event may occur. It would seem to exclude a certainty of belief that it will occur. Thus we can be said to have no risk of death; the risk is only as to time, place and manner.” (Seavey, 1927, p. 7)

the range of properties that may trigger the principle, irrespective of the attendant risks: **First**, a danger-making property cannot itself be the property of being harmed by the act or a property that is constitutive of such harm (such as the property of suffering contusions as a result of this act). Absent this constraint, *Danger Zone* would imply that the railroad did violate the duty owed to Helen Palsgraf, insofar as it was foreseeable to the guard that someone or other would be harmed as a result of the push. **Second**, if a risk of harm is associated with a certain property only because having that property makes it more likely that one also has some other property — more exactly, some other property which is itself danger-making (that is, which satisfies the antecedent of *Danger Zone*, including the constraints) — then the former property is not itself danger-making. For example, if a danger-making property, relative to a certain action, is the property of being on a certain basketball court at noon tomorrow, then the mere fact that belonging to a certain club makes it very likely that one will be on the basketball court at that time does not make club membership a danger-making property in the relevant sense. If the negligent act unforeseeably injures a club member in a distant location, the negligent actor has not violated her rights. Of course, I do not mean to suggest that no more than one property can serve as danger-making relative to a given act — for example, a certain act may pose distinct risks for those within earshot and those within eyeshot, or to those who are pregnant and those who are not. **Third**, a property is danger-making relative to some act only if it is foreseeable to the agent that her act carries an *appreciable* risk of harming *whoever* instantiates the property. Note that an appreciable risk need not be substantial. In a fair lottery, each ticket-holder has an appreciable chance of winning, even if the odds are extremely low. By contrast, in ordinary circumstances the theoretical possibility that by turning on my desk lamp I will electrocute my next-door neighbor constitutes an unappreciable risk. Given this third constraint, an act satisfies

the antecedent of *Danger Zone* only if it both creates an unreasonable risk of harming *someone or other* who instantiates the danger-making property and an appreciable risk of harming *whoever* instantiates that same property.¹² (Since this constraint is not put to work in this paper, I will leave it as an exercise for the interested reader to discern its motivation.)¹³

Finally, a bit of terminology: let us define the term “danger zone” in terms of danger-making properties. All and only those who instantiate the danger-making property during the relevant time interval are occupants of the danger zone. *Danger Zone* says that by performing a negligent act, I violate a duty of care owed to all and only those who occupy the danger zone. Additionally, throughout this paper I use the term “negligent act” to refer to any act that satisfies the antecedent of *Danger Zone*, irrespective of whether anyone winds up in the danger zone.

Danger Zone has four important implications that I wish to highlight; indeed, it is these implications that lend substantial support to the attribution of the conception to Cardozo and to tort law more broadly. The first implication concerns “lottery-like” cases. Just as I may

¹² Strictly speaking we should also build into the constraint the further condition that the two risks (that is, the appreciable one and the unreasonable one) are owing to the same foreseeable process eventuating in harm.

¹³ A fifth and final point of clarification: While “unforeseeability comes in four categories: (i) unforeseeable plaintiffs, (ii) unforeseeable types of harm, (iii) unforeseeable extent of harm, and (iv) unforeseeable manner of harm,” *Danger Zone* concerns only the first (Smith 2011, 23).

Whereas the last two categories do not generally shield the negligent actor from liability, some authorities have recognized category (ii) as excusing; however, this result is reached on the basis of a proximate-cause requirement (rather than breach-of-duty). (Smith, 2011, pp. 19–20, 22–25)

reasonably believe both that a certain lottery will have a winner and that any given ticket-holder has extremely low odds of winning, so too I may reasonably believe both that my act creates an undue risk of harming someone or other in the danger zone and that it exposes no individual to more than an extremely low risk of suffering harm. In such a case, if there are occupants of the danger zone, then *Danger Zone* implies that I have wronged them.

The second and third implications can be brought out by considering the following hypothetical: today I perform an act that creates an unreasonable risk of harming anyone who will be in a certain building tomorrow at noon. At the time of action, I have decisive reason to believe that at noon tomorrow Hector will be in the building and Bertha will be on the other side of town, out of harm's way. As it happens, the evidence is doubly misleading, and it is Bertha rather than Hector who winds up in the danger zone. *Danger Zone* implies both that I have wronged Bertha and that I have not wronged Hector, notwithstanding the antecedent risks. Nobody who has any feel for the common law would doubt these conclusions. If the risk of harm materializes, and Bertha suffers injuries as a result of my act, I cannot respond to a demand for compensation with the assertion that "I couldn't have known that you'd be there."¹⁴ Likewise, if my act unforeseeably injures Hector on the other side of town, he cannot say, by way of answering my invocation of *Palsgraf*, "but you thought I'd be in the building."¹⁵

The fourth and final implication relates to the fact that there are two ways in which it might be unforeseeable to a negligent actor that a certain individual, X, will wind up in the danger zone. First, the actor may have reason to assign a very high probability to the proposition

¹⁴ Weinrib (2003, p. 165); Schwartz (1999, n. 71).

¹⁵ Goldberg and Zipursky (2021, p. 200).

that X will not occupy the danger zone. Second, the actor may lack any acquaintance with X, such that he may be unable to form any beliefs about X, including the belief that X will wind up in the danger zone. *Danger Zone* is not sensitive to this distinction, as its conditions do not require that the agent form beliefs about the particular individuals who wind up in the danger zone.

In each of these four respects, *Danger Zone* contrast sharply with the following alternative rule, which (on plausible factual assumptions) would also support Cardozo's denial of relief to Helen Palsgraf:

*Naïve*_(me): for all persons X, if I believe or should believe that by ϕ -ing I would create an unreasonable risk of harming X, then I owe it to X not to ϕ .

As I have said, *Naïve* yields the opposite result in all four cases. For example, given my justified belief that Bertha was not in harm's way, *Naïve* implies that I did not wrong her. Similarly, if a given individual is in no sense on my radar, such that I cannot form beliefs about them, then *Naïve* implies that I owe them no duties of care.¹⁶

¹⁶ To avoid this implication (as well as the one about Bertha), it would be natural to modify *Naïve* by appealing to danger-making properties. On this modified version, if an act is unreasonable in virtue of the risk it imposes on any arbitrary danger-zone occupant, then the agent owes it to *every* danger-zone occupant not to perform the act. However, this modified version reaches the wrong result in lottery-like cases (more on which below). Moreover, it is vulnerable to exactly the same challenge that will be raised for the rights-based account of *Danger Zone* (involving "parity constraints"). Something akin to this modification of *Naïve* is

Before commenting further on *Danger Zone*, let us introduce the final competing conception, which will serve as its chief foil.

Andrews' Rule: If an agent should believe that a certain act creates an unreasonable risk of causing harm (to someone or other), then that agent owes it to every person not to perform that act.

Although I do not go so far as to attribute this conception to Andrews' dissenting opinion, it not only accords with his bottom line, but is also suggested by his arresting remark that a negligent actor commits "a wrong not only to those who happen to be within the radius of danger but to all who might have been there—a wrong to [everyone]."

II Interest Theories

A. Duty of Care

Let us now turn to theories of "directed duties" and "claim-rights" with an eye to identifying the relational notions at work in the above conceptions of the duty of care. According to the family of views that we will first consider — interest theories, broadly construed — to call an obligation directed toward a certain individual is just to make a claim about how the individual's interests figure in an account of the obligation's grounds.¹⁷ According to the traditional Interest Theory of rights, widely associated (nowadays) with the work of Joseph Raz, "to say that a person has a right is to say that an interest of his is sufficient ground for holding another to be subject to a duty, i.e., a duty to take some action which will serve that interest or a duty the very

put forward in Hirsch (forthcoming); I am greatly indebted to Hirsch's paper, which is marred only by its failure to consider *Danger Zone*.

¹⁷ For discussion, see Thompson (2004, pp. 348–49).

existence of which serves such interest” (Raz, 1984, p. 5). Applying this general theory to the law, Raz maintains that an individual has a legal right if “the law holds his interest to be sufficient ground to hold another to be subject to a duty” (*id.*, p. 14). On the other side of the ledger, “a duty is towards a certain person if and only if it is derived from his right” (*id.*, n. 14).

Proponents of *Danger Zone* can enlist the Interest Theory only if the duty not to act negligently is founded on the interests of all and only those who (actually) occupy the danger zone. Such a claim faces insurmountable difficulties. Let us put forward the following principle governing the grounding of duties on interests:

Parity Constraint: If a certain course of conduct, as well as an obligation requiring or forbidding that conduct, have indistinguishable effects — both actual effects and expected effects — on the interests of two individuals, then the obligation to perform, or refrain from performing, that course of conduct is grounded in the interests of one of the individuals only if it is also grounded in the interests of the other.

To be sure, the Parity Constraint is subject to the proviso that the individuals cannot be distinguished by appeal to special relationships, social roles, or prior interactions, each of which may require or entitle us to be more solicitous of the interests of some individuals than of others. More generally, it is subject to the proviso that if other conditions, beyond those concerning the actual and expected effects on an individual’s interests, are necessary to generate the relevant duty, then these conditions do not serve to distinguish the individuals. In light of these provisos,

the Parity Constraint can hardly be called into question, amounting to a mere application of the more general requirement that “like cases be treated alike.”¹⁸

Given the Parity Constraint, it would appear that *Danger Zone* cannot enlist the support of the Interest Theory. It is not merely possible but positively routine for two individuals, with equal ex-ante probabilities of danger zone occupancy, to find themselves on opposite sides of the boundary marking the danger zone. Indeed, the individual who winds up outside the danger zone may have had the higher ex-ante chance of occupying it. As for actual effects, the individuals may be either equally affected or equally unaffected by the negligent act. For example, if the negligent activity involved a failure to properly dispose of combustible materials, an ensuing explosion may have been unforeseeably powerful (injuring everyone within a fifteen-block radius, say) or may not have materialized at all.

This argument rests on an important assumption. While people have interests in not being harmed — alternatively, while people have interests the setting back of which constitutes harm — they do not, in general, have an interest in not being harmed in the danger zone. (Of course, this may not be true of particular danger zones, such as one’s place of worship or family home, but the general point stands.) Similarly, they do not have an interest in not *occupying* the danger zone; this becomes especially clear when it is borne in mind, first, that by the time an individual winds up in the danger zone the probabilities that any danger zone occupant will be harmed may have fallen considerably; second, that the probability that determines whether a given property is danger-making is fixed by the evidence available to the agent rather than the occupant. (And

¹⁸ Of course, the Parity Constraint is “merely” a normative requirement for legal systems, which is all that my argument requires.

even if the probability *were* assigned from the occupant's standpoint, it would be a mistake to equate danger-zone occupancy with the property of, e.g., being afraid.) The construction of a danger zone is relevant for the determination of whether an act poses an unreasonable risk of harm, as well as for the assignment of probabilities of injury to particular individuals — for example, if my act carries a 10% chance of injuring anyone inside a certain building tomorrow at noon (and to nobody else), and if a certain individual has a 10% chance of being in the building at that time, then my act carries a 1% chance of serious injury to that individual (assuming independent probabilities). As far as that individual's personal interests are concerned, this is the only probability that counts in evaluating my behavior.

Perhaps one will dispute this and say that the antecedent odds of injuring the actual danger-zone occupant (e.g., Bertha) should be identified with the antecedent odds of injuring a danger zone occupant (whoever she is), regardless of whether the actual occupant (Bertha) had low odds of winding up in the danger zone. In a moment I will show that this proposal, even if correct, would not succeed in reconciling the Interest Theory and *Danger Zone*. However, we should also reject the proposal. The interest in avoiding physical injury is, of course, the interest of an individual, independent of how she is designated; likewise, the ex-ante odds of injury that are at issue are the odds of injuring particular individuals. To be sure, it is a delicate question which of the many ways of designating an individual should be used to determine the epistemic probability that a certain individual will suffer harm. But this much is certain: when the only basis for saying that harm to a certain individual was foreseeable to the negligent actor is that the individual came to occupy the danger zone, then in determining that individual's ex-ante chances of getting harmed we cannot ignore the agent's reasons for believing that the individual would not occupy the danger zone. Just as we would not say of a lottery winner that the fact that their

number was drawn shows that their chances of winning were not low to begin with, so too with the individual who unforeseeably winds up in the danger zone. In short, we cannot identify the antecedent probability that a certain individual would suffer harm as a result of an act with the antecedent probability that they would suffer harm *if* they wind up in the danger zone.¹⁹

Even if the proposal were correct, however, it would not serve to reconcile *Danger Zone* with the Interest Theory. To see this, consider a negligent act with the following two features: first, the danger zone is densely populated, such that the negligent actor has reason to believe both that an act carries an unreasonable risk of harming someone or other in the danger zone and that it carries an extremely slight (yet appreciable) risk of harming any given danger zone occupant. Second, the act creates the same risk of harm for those who instantiate the danger zone and for at least some of those who do not. For example, suppose that an act is negligent because it risks seriously harming a very small number of people in a densely-populated danger zone, such as Times Square on New Years Eve. Suppose further that the act also has a very small chance of harming someone in a certain one-bedroom apartment unit outside of Times Square, in which event the occupants of Times Square would be spared. Even though the act (let us suppose) is not negligent on account of the latter risk, the antecedent risk of harming any given reveler in Times Square may be the same as the risk to the occupant of the apartment, given the different sizes of the respective populations. In nevertheless distinguishing between occupants and non-occupants of the danger zone, *Danger Zone* violates the Parity Constraint, even if an individual's ex-ante risk is evaluated in accordance with the proposal under consideration.

¹⁹ I will later show that the “ex-ante” interest theorist has additional reason for rejecting the proposal (see n. 27).

A second objection to my argument concerning the incompatibility of *Danger Zone* and the Interest Theory concerns the relevance of the second proviso for legal rights in particular. One might think that certain practical considerations justify conditioning a duty not to perform an unreasonably risky act on factors other than the interests of an individual who would be unreasonably endangered, and that among such conditions is that the endangered individual winds up in the danger zone. If this were correct, then the second proviso would not be met, and the Parity Constraint not violated after all.

What kind of practical considerations would justify conditioning the duty on the materialization of the risk that the danger zone would be occupied? Since even decent humans can be counted on to act negligently from time to time, and decent courts to mistake reasonable acts for unreasonable ones, unrestricted liability for negligently-inflicted harm is haunted by the specter of far-reaching and indeterminate liability that would, it has been argued, dampen the entrepreneurial spirit and give rise to unjustifiable administrative costs. Even if such costs were significant, however, it is important to see that they would not provide any reason for conditioning the *duty* not to act unreasonably on danger-zone occupancy. The Interest Theory under consideration distinguishes sharply between legal duties and liability for damages. And once these notions are prized apart, it is easy to see that these costs favor restrictions on liability alone.

A second reply to the objection is also available, though only for those who think (with Cardozo and Andrews) that the duty of care may be violated absent the infliction of harm.²⁰

²⁰ Here I am relying on the assumption that a duty of care can be violated absent harm, but the argument merely supplements the previous one.

Consider two individuals with very slight antecedent risk of danger zone occupancy — let us suppose that one of them (despite the low odds) winds up in the danger zone, but that neither of them ultimately suffers harm. With respect to such individuals, how can we justify the differential treatment by appealing to “other conditions” beyond those which concern the actual and expected effects on their interests? Given the negligible risk, neither individual satisfies the condition that *does* concern the impingement of the required act, or of the duty itself, on their interests. And if neither individual satisfies the interest condition, then the duty at issue (however well grounded) is not grounded in their interests and so, according to the Interest Theory, cannot be said to violate their rights. To bring the point home, suppose that by the time an individual comes to occupy the danger zone, the chances that danger-zone occupants would suffer harm has fallen precipitously. At no time did such an individual face a substantial risk of suffering harm as a result of the negligent act; should we nevertheless say that the duty not to perform such an act was grounded in her interests?²¹

²¹ An alternative theory of the condition’s rationale takes off from the observation that if an individual outside the danger zone suffers harm, then such harm is not the outcome of a process the foreseeability of which rendered the action unreasonably risky — in the legal vernacular, the resulting harm would not be “within the risk” in virtue of which the act was unreasonable. Given this fact, it is a natural thought that a more general “harm within the risk” condition is what explains why the interests of those who wind up outside the danger zone do not ground a duty that the agent violates when he performs the negligent act. There are at least three problems with this suggestion: First, if (actual) harm is not required to violate the duty not to perform the unreasonably risky act, then neither is harm that falls “within the risk”. Second, even if harm

The third and final objection I will consider concerns the scope and generality of the obligation recognized by *Danger Zone*, which (let us stipulate) provides both necessary and sufficient conditions for relational duties of care with respect to unintentional harms. In particular, proponents of *Danger Zone* do not recognize a more general conditional (directed) duty of care, from which the unconditional duty articulated in *Danger Zone*'s consequent can be derived. Just like, on a traditional view, the directed duty incurred by a particular promise does not derive from an antecedent duty, owed to the promisee, that I keep any promises that I make to her, so too with *Danger Zone*'s conception of the duty of care. The third objection takes issue with this feature of *Danger Zone*, maintaining that the duty recognized by Cardozo should be construed as conditional in content and that such a construal eliminates the challenge owing to the Parity Constraint. In particular, the objector puts forward the following alternative principle as the principle underlying the *Palsgraf* decision:

were a condition of violating a duty of care, we have already seen (n. 13) that it is no condition of legal liability that the harm arise in a foreseeable *manner*, which is enough to show that there is not a general “harm within the risk” condition of liability as a matter of positive law. Third, whether or not a “harm within the risk” condition is or should be a condition of *liability*, why insist on it as a condition of the *duty* not to perform the unreasonably risky act? It is not enough to say that the condition requiring danger-zone occupancy for breach of duty follows from a more general “harm within the risk” condition if the latter condition itself lacks justification (for all the same reasons as the former).

DZ*: For every pair of individuals, each member of the pair is under a standing obligation not to perform a negligent act, provided that the other member of the pair winds up in the danger zone (relative to that act).

According to **DZ***, each of us has as many standing obligations as there are other people; in my own case, each such obligation requires me not to perform a negligent act, provided that a certain person (one for each standing obligation) winds up in the danger zone. Now, when Bertha winds up in the danger zone against all odds, **DZ*** agrees with *Danger Zone* that I fall under an unconditional obligation, owed to Bertha, not to perform the negligent act. However, the proponent of **DZ*** would say that this unconditional obligation is merely derived from the conditional one (also owed to Bertha), which requires that I not perform a negligent act if Bertha winds up in the danger zone. Given that the unconditional obligation is derived in this way, it does not implicate the Parity Constraint, since (in contrast with the more general conditional obligation) it is grounded not in Bertha's interests, but rather in the more general conditional obligation.²²

This objection appeals to standing conditional obligations that are directly grounded in the interests of the individual who figures in the condition. However, it strains credulity to suppose that Bertha's personal interest in avoiding physical harm directly grounds the conditional duty (call it "**DZ***_(me, Bertha)") that I not perform a negligent act if Bertha winds up in the danger zone. To see this, let us consider the possibility that Bertha's personal interests ground the following close cousin of *Naïve*:

²² Note that while the conditional duty is grounded in Bertha's interest, it does not violate the Parity Constraint.

Less Naïve (*me, Bertha*): I am obligated to refrain from performing any act ϕ , provided that there is a property F with the following two features: a) the risk of suffering harm that my ϕ -ing creates for *anyone* who instantiates F, together with the probability that F will be instantiated by *someone or other*, are sufficient to render my ϕ -ing unreasonable; and b) Bertha instantiates F.

*Less Naïve*_(*me, Bertha*) diverges from *DZ**_(*me, Bertha*) in ways that bring into relief the problems with the suggestion that the latter is directly grounded in Bertha's interests. Most significantly, unlike the latter, the former (like its cousin *Naïve*) does not prohibit negligent acts in those lottery-like cases where each danger zone occupant (whoever she is) faces extremely low odds of injury. Although this implication of *Less Naïve*_(*me, Bertha*) prevents us from attributing it to Cardozo (and to tort law), it would seem to be a necessary feature of any duty grounded directly in Bertha's interest: for how can Bertha's interests directly ground a duty forbidding me from imposing only a trivial (albeit appreciable) risk of harming her, even on the assumption that she is a danger-zone occupant? And this appears to show that even if Bertha's interests ground *Less Naïve*_(*me, Bertha*), they cannot ground *DZ**_(*me, Bertha*).

However, this reply to the objection is arguably too hasty. Even if Bertha's risk of injury is very low, the objector might argue that her interests nonetheless ground a duty not to impose this risk in the absence of competing factors pulling the other way; and given that the act in question poses an unreasonable risk of harming someone or other, the act's bearing on Bertha's interests may be sufficient to render the act unreasonable.

Even if this argument were valid, it would not follow that Bertha's interests ground *DZ**_(*me, Bertha*); rather, the argument would support the view that those interests ground an

alternative rule, one that (despite its applicability to lottery-like cases) is flatly inconsistent with the decision in *Palsgraf*:

*DZ***_(me,Bertha): I am obligated not to perform any negligent act, provided that such an act harms Bertha.

While this rule contains a “fact-relative” element, so does *Less Naïve*: Just as it may have been unforeseeable to the negligent actor that Bertha would be harmed, so too it may have been unforeseeable that Bertha would wind up in the danger zone. Moreover, *DZ***_(me,Bertha) hews more closely to Bertha’s interests than does either *Less Naïve*_(me,Bertha) or *DZ**_(me,Bertha), given that Bertha both has an interest in not being harmed even if she is outside the danger zone and lacks an interest in not occupying the danger zone. Accordingly, the third objection fails, since it is not plausible that the standing obligations in question (instances of the *DZ** schema) are grounded in the interests of the corresponding individuals (e.g., Bertha).

Before moving on from the traditional Interest Theory, let us briefly consider the relation between that theory and *Andrews’ Rule*. The first thing to note is that the Parity Constraint does not prevent us from assigning the Interest Theory to *Andrews’ Rule*: in assigning rights to everyone, *Andrews’ Rule* does not distinguish between individuals who are subject to the same risks of suffering the same harm. Nevertheless, one may harbor doubts as to whether the personal interests of, e.g., a far-flung lighthouse keeper ground my obligation not to drive recklessly on a busy city road. After all, we may suppose that there is no sense in which the lighthouse keeper’s life would have gone better had I not cut-off the slow city driver, or gone worse had I not been

obligated to stay in my lane.²³ However, insofar as an interest theory can appeal to the conditional obligation *DZ***, then even if the Interest Theory and *Andrews' Rule* are incompatible, the interest theorist can nonetheless reach the conclusion that the negligent guard violated the rights of Helen Palsgraf.²⁴

²³ Raz does allow that certain individual rights can be “justified [i.e., grounded] ultimately” by appeal to the interest of the general public rather than the right-holder — e.g., a journalist’s right to protect her sources is grounded ultimately not in that journalist’s own good (well-being), but rather in the usefulness to the public of her “being able to collect information” (1986, p. 179). Correspondingly, might not the lighthouse keeper have a right to my conscientious driving that is grounded in the interest of the public? Although I welcome the proposal, which would supplement the justification of *Andrews' Rule* that I shall suggest later in the paper, the cases are not perfectly analogous. For Raz explicitly affirms that any given journalist has a personal interest in “being able to collect information,” merely denying that their corresponding right is *non-instrumentally* grounded in that interest (*ibid.*). By contrast, in the case of the lighthouse keeper it is a challenge to articulate a personal interest that might even *instrumentally* ground the relevant right.

²⁴ I note in passing that the foregoing analysis casts doubt on Nicolas Cornell’s argument (2015) that if Helen Palsgraf was properly aggrieved (as he controversially intuits), then this shows that one can be properly aggrieved by conduct that does not violate one’s rights. In Cornell’s view, while Cardozo was correct in holding that the plaintiff’s rights were not violated, he was wrong to think that being properly aggrieved requires the violation of one’s rights. However, as we have just seen, an Interest Theory of rights can uphold the distinction between rights-violation and

Let us turn to a contractualist variant of the Interest Theory of claim-rights, one that applies specifically to the moral domain. (Although the next section contains a broader discussion of moral claim-rights, it will be useful to first introduce contractualism in connection with the discussion of *Palsgraf* and *Danger Zone*.) Speaking very generally, contractualism is a moral theory that sets out to resolve conflicts of interest in a manner reflective both of the *separateness* and the *equal significance* of every person whose interests stand to be affected by a given course of conduct. Oversimplifying somewhat, let us say that the contractualist holds an act to be morally impermissible when the evidence available to the agent suggests that it will set back *somebody's* interests to a greater extent than would some available alternative set back the interests of *anyone*. In other words, an act is impermissible when the strongest individual objection to it (based on the objector's personal interests) is greater than the strongest individual objection to an available alternative.²⁵ R. Jay Wallace (2019) has recently extended

proper aggrievement and yet still deliver the result that Helen Palsgraf's rights were indeed violated. And as we will see below, the same is true of the Quality of Will conception of the duty of care: it too is consistent with the view that Helen Palsgraf's rights were violated. Cornell's own theory of Proper Aggrievement has been criticized by Wallace (2019, pp. 195–96) and Ramakrishnan (m.s.).

²⁵ This is an oversimplification of the contractualist view — most glaringly, it fails to make explicit the role played by principles in accounting for the impermissibility of an act (Scanlon, 1998). The omitted complication only creates further difficulties for Wallace's theory of claim rights, as he recognizes (2019, ch. 5, n. 46; cf. Jonker, 2019). On independent grounds, Liam Murphy (2021) has recently argued that so-called "act contractualism," which drops the appeal to

contractualism to produce an account not only of morally permissible action but also of moral rights. When a certain act would be impermissible by contractualist lights, each person with an objection strong enough to account for the act's impermissibility holds a claim-right against the agent not to perform the act. Although Wallace (2019, pp. 196-98) explicitly maintains that a contractualist theory of rights is consistent with the majority ruling in *Palsgraf*, it is in fact inconsistent, and for the same reason as the traditional Interest Theory. That is, it cannot explain the differential treatment accorded to two individuals, subject to the same probabilities of suffering harm, who end up on opposite sides of the danger zone.

This incompatibility between *Danger Zone* and the contractualist theory of claim rights is independent of Wallace's "*ex-ante*" interpretation of contractualism (EAC). According to EAC, the full extent of any given individual's personal objection to a given act is based on the act's expected effects on *that* individual's personal interests. In particular, in evaluating the permissibility of courses of conduct, EAC does not factor in the antecedent probability that *someone or other* would be harmed by the act; rather, the evaluation only takes into account the expected impact of the act on each individual's personal interests. The attentive reader may have already noticed that EAC is a version of *Naïve* (with "unreasonableness" receiving a contractualist gloss). In the previous paragraph I made the point that the contractualist theory of rights (like all interest theories) cannot account for *Danger Zone*'s allocation of rights. The present point is that, since Wallace rejects *Danger Zone* outright (in favor of *Naïve*), his view is

general principles in the account of moral permissibility, is to be preferred over Scanlon's original version.

only reconcilable with *Palsgraf* if *Naïve* can be imputed to Cardozo — an imputation we have already ruled out.

More importantly, many of the problems that rendered such an imputation implausible also count against EAC as a theory of moral claim-rights. Consider the following lottery-like case: in order to save my marriage I engage in unreasonably risky behavior — for example, to avoid missing my early-morning train, I knowingly fail to mop up the liquid that I spilled on the smooth pink marble stairs in Grand Central Station. We may suppose that my conduct is negligent, since it is sufficiently likely that someone or other will be seriously injured walking down the stairs. We may also suppose that, given the evidence available to me, any given individual's odds of suffering harm are very low — this is Grand Central Station after all (and, given the odd hour, nobody is in the immediate vicinity at the time of the spill). Finally, since saving a marriage is no small thing, we may suppose that the act would not have been unreasonable if the risk of causing harm (to someone or other) were as small as the risk imposed on any given person. Given these facts, Wallace cannot account for the fact that if the risk materializes and an individual slips on the stairs, then we would expect all of the rituals associated with relational morality to apply: resentment, apology, repair, forgiveness...the whole dance. For Wallace is committed to holding that I did not violate (or otherwise manifest disregard for) a claim of the injured party — a condition, in his view, of proper aggrievement.²⁶

²⁶ This case is similar in structure to Villain 2 discussed in Horton (2017), though Horton is not concerned with the relational dimension. Moreover, my case (in contrast with Horton's) is not vulnerable to Wallace's charges (issued against Frick, 2015) of being artificial or of illicitly relying on a bureaucratic or administrative context (Wallace, 2019, p. 227). It is also worth

And the counterintuitive implications are not limited to lottery-like cases. In my opening hypothetical, for example, Wallace is committed to the view that Bertha has not been wronged, given her trivial antecedent risk of suffering harm.²⁷ Given these implications, Wallace's theory belies his (non-revisionist) intentions.

B. Morality of Harm-Doing

Let us continue extending the foregoing analysis of *Palsgraf* to the *moral* right not to be (unjustifiably) harmed; in addition to considering other interest theories of moral rights (beyond

emphasizing that even if an ex-ante contractualist were to supplement their theory to account for the wrongfulness of my conduct in *Grand Central* (e.g., in the manner of Frick, 2015), it would not follow that they have accounted for the relational dimension — *in particular, for the fact that the person who slipped is properly aggrieved.*

²⁷ Earlier, I considered the proposal that the antecedent odds of injuring the actual danger zone occupant should be identified with the antecedent odds of injuring a danger zone occupant (whoever she is). As noted (n. 19), the ex-ante contractualist has special reason to reject the proposal: namely, it is inconsistent with Frick's thought experiment, Mass Vaccination (*Unknown Victims*), which has served as the primary motivation for ex-ante contractualism (Frick, 2015, p. 181). In Frick's example, having "gene G" is a danger-making property, and all and only those children with the gene (when the vaccine is administered) are in the danger zone. If the ex-ante contractualist were to adopt the proposal under consideration, then they would have to say that, for each child who has gene G, the ex-ante probability of death is a near certainty, in which case the intuition that administering the vaccine is permissible would not be explained. Cf. Setiya (2023, pp. 319–21).

EAC), we will also extend the analysis to embrace the intentional infliction of harm. Before proceeding, it will help to clarify the relevance of such moral rights (i.e., directed obligations) for our overarching question concerning the conditions of proper aggrievement. According to one family of views, an individual is properly aggrieved by another's conduct only if the conduct violates an obligation that stands in a certain relation to the individual's interests — the same relation in virtue of which the obligation is owed to that individual. In the remainder of this section, I will consider three such interest theories — two evidence-relative theories as well as a fact-relative theory — and demonstrate that they each fail to account for the phenomena. In the final part of the paper I turn to another family of views, united by the thesis that an individual is properly aggrieved by another's conduct if the conduct manifested insufficient concern for the individual's interests or rights. It bears noting that the two families can be combined in various ways: for example, some views in the first family hold that the manifestation of insufficient concern for an individual's claim-rights constitutes another necessary condition for proper aggrievement, over and above the infringement of the claim. Additionally, a certain kind of interest theorist might hold that any act that manifests insufficient concern for another person violates the latter's claim for sufficient concern.

To focus our discussion of evidence-relative interest theories — that is, theories which assign claims on the basis of expected outcomes on individual interests, as determined by the evidence available to the agent at the time of action — I will consider contractualist theories of claim-rights. (The arguments straightforwardly generalize to non-contractualist evidence-relative interest theories.) We need not give EAC further consideration: I have already shown that EAC generates highly revisionist results in cases of negligent conduct, and nothing changes when we tweak the above examples by raising the agent's credences to the level of "substantial certainty".

Would an evidence-relative ex-post contractualist theory of claim rights (“EPC”) perform any better? As a first pass, let us formulate EPC as follows: where an agent should believe that their act would unjustifiably harm someone or other — that is, would unjustifiably set back their personal interests — then irrespective of the odds of harming any given individual, the act violates the claim of whoever is in fact harmed. To demonstrate that EPC fares no better, I will now consider three versions of a case involving the following common features: in order to save my job, I knowingly cause an explosion in a certain field; I justifiably believe both that Lutz will be on the other side of town at the time of the blast and that anyone on the other side of town will be out of harm’s way; as it happens, while the former belief is true, the latter belief is false, and the unforeseeably powerful explosion injures Lutz on the other side of town; finally, let us stipulate that the act has no normatively significant features beyond my continued employment and the possibility that someone will suffer direct harm from the blast. Notwithstanding these common elements, the cases differ as follows:

Version 1: As a result of performing due diligence, I knew that nobody would be on the field at the time of the blast, and justifiably believed (falsely) that nobody would be injured by the blast.

Version 2: I justifiably believed that some (unidentified) individual would be on the field and would be injured by the blast. As it turned out, the evidence was doubly misleading; nobody was on the field (and Lutz was injured on the other side of town).

Version 3: Same as version 2 except that somebody (a certain Mutz) was on the field and was injured. The ex-ante odds that Mutz in particular would be on the field were as low as the ex-ante odds that Lutz would be on the field.

In Version 1, EPC, like any other plausible evidence-relative theory, delivers the fairly intuitive result that I did not wrong Lutz, notwithstanding the injury. (If feelings of guilt would be natural, evidence-relative theories would chalk this up to the more general phenomenon of so-called “agent regret”.) In Version 2, even if my act was blameworthy, no evidence-relative Interest Theory of rights can consistently dispute the same conclusion — namely, that Lutz was not wronged.²⁸ For in addition to the *interpersonal* Parity Constraint considered above, any Interest Theory of claim-rights is also subject to an *intrapersonal* Parity Constraint. I will begin with an unqualified statement of the latter constraint, before adding a proviso that does not bear on the present discussion. If two acts by a given agent (as well as obligations to refrain from performing each act) have the same actual and expected effects on the personal interests of a certain individual, Z, as well as on the personal interests of the agent, then Z’s interests ground an obligation not to perform one of the acts only if they ground an obligation not to perform the other. As noted, this formulation requires qualification.²⁹ The need for qualification is based on the (characteristically contractualist) assumption that if an act is reasonable, then it does not violate anybody’s claims. Given this (controversial) assumption, we must add to the intrapersonal Parity Constraint the proviso that neither of the acts has (actual or expected) effects

²⁸ I make no claims about the law’s treatment of Version 2. In particular, we need not concern ourselves with the doctrine of “transferred intent” (applicable to intentional torts). For a discussion of whether the *Palsgraf* holding can be reconciled with the latter doctrine, see Ramakrishnan (m.s.).

²⁹ This proviso can be placed under the rubric of the same general proviso that qualified the interpersonal parity constraint.

on the interests of third parties that render the act reasonable (where it would have otherwise been unreasonable). After all, if Z demands that the agent not perform a certain act due to its expected effects on his personal interests, the act's expected effects on third parties may enable the agent to silence Z's demands as follows: "Z, I know that it will be bad for you if I perform this act, but it would be at least as bad for others if I refrain, and who is to say that your blood is redder than theirs?" As noted, the proviso does not bear on the case of Lutz; for in none of the three versions does the act's expected effects on third parties contribute to making the act reasonable (where it would otherwise have been unreasonable). Accordingly, we conclude that a proponent of EPC must hold that since Lutz's rights were not violated in Version 1, they were not violated in Version 2 either.³⁰

Can EPC account for this judgment regarding Version 2 – that is, the judgment that while my act may have been blameworthy, Lutz had no personal claim against me not to perform it? Not according to the above formulation of EPC, which made no mention of danger zones. (After all, I had reason to believe that somebody would be harmed, and somebody was indeed harmed!) So let us (charitably) modify EPC accordingly. Rather than recite the conditions of EPC2, let us

³⁰ To be maximally concessive to the interest theorist, one might consider drawing an additional proviso excluding cases where one of the acts violates Z's rights due only to its combined (aggregative) expected effects on Z's interests and those of third parties. Such a proviso would not help the proponent of EPC: since (let us stipulate) it was neither foreseeable (appreciable) that Lutz would be injured nor that anyone in his position would be injured, Lutz has no personal interests that can (on an evidence-relative approach) piggyback on, or join forces with, those of the third parties to ground a claim.

stipulate that it is exactly like *Danger Zone* except that the notion of “risk” is no longer to be understood to exclude “substantial certainty,” while the notion of “unreasonableness” is to be given a contractualist interpretation.³¹ EPC2 reaches the correct verdict that Lutz has no claim against us; after all, he never made it to the danger zone. When we turn to Version 3, EPC2 yields the result that I wronged Mutz but not Lutz. While I think this result accords with intuition, EPC is not entitled to it, as it violates the original (interpersonal) Parity Constraint: given that the actual and expected effects on their interests are the same, interest theories (including EPC) must hold that if the blast wrongs Mutz, it must also wrong Lutz. In short, while proponents of EPC would not deny that Mutz’s rights *were* violated in Version 3 or that Lutz’s rights were *not* violated in Version 1, this combination of views is unavailable to them in light of the parity constraints. Finally, I note that nothing in the foregoing analysis changes if we reduce the antecedent risk of injury (in Versions 2 and 3) to the level where the wrongful act is merely negligent — that is, the argument straightforwardly applies to the morality of negligent and intentional harm-doing alike.

Before setting aside the Interest Theory, let us briefly consider a fact-relative theory of claim-rights, such as Thomson’s “Harm Thesis”, according to which an agent infringes the claim-right of anyone they harm, irrespective of whether the agent was at fault (Thomson, 1992, ch. 9). (The theory is “fact-relative” in that its assignment of claims is a function of actual effects rather than expected ones.) In order to bring it closer to the previously-considered evidence-relative views, let us modify Thomson’s Harm Thesis in one respect: instead of saying that a

³¹ Additionally, whereas *Danger Zone* concerns what an agent “believes *or* should believe,” EPC considers only the latter.

harmful act always infringes upon the claim of the injured party, let us say that the harmful act infringes upon the latter's claim unless its other consequences or features are such as to justify the harm, irrespective of whether the agent had reason to believe that the justifying facts would obtain. Since this view implies that I infringed upon Lutz's claim in Version 1, it is not vulnerable to the objection based on the parity constraints. Nevertheless, whatever its merits, this view creates a chasm between the infringement of a claim and proper aggrievement.³² Thomson herself observes that when an agent inflicts harm faultlessly, subsequent remorse or guilt would not be appropriate, given that "what you feel is not properly described as 'regret' or 'guilt' unless you think you were at fault."³³ Plainly, what is true of remorse and guilt is also true of resentment, apology, and forgiveness.³⁴ Accordingly, if a fact-relative theory of claim-rights is

³² As Thomson (1992, p. 229) observes, if it was not foreseeable to me that my flipping of a light switch in my home would electrocute my neighbor, then I am not at fault for causing their injuries despite the violation of their claim. Conversely, if I attempt to harm someone without justification, they would be properly aggrieved even if I were to miss the mark.

³³ *Id.*, 97. Thomson does allow that so-called "agent regret" is possible here, but she denies that this is owing to the infringement of claims (*Id.*, pp. 241–42).

³⁴ As for the duty of repair, matters are slightly more complicated. On the one hand, in an earlier chapter Thomson argued for the thesis that even a faultless breach of a promise infringes a claim-right partly on the ground that this would explain the promisor's duty to compensate the promisee for resultant harms, notwithstanding the absence of fault (*id.*, 96). On the other hand, mention of such a duty to compensate is conspicuously absent when Thomson later argues (in ch. 9) for the parallel conclusion that even a faultless infliction of harm infringes the sufferer's

correct, then the relation between claim-rights and proper aggrievement is at best highly attenuated. A natural response from a proponent of such a theory is to adopt either of two strategies: the first strategy involves supplementing their fact-relative theory of rights with a belief- or evidence-relative one (corresponding with the “subjective ought”), and linking proper aggrievement to the rights recognized by the latter theory. Such a strategy would, of course, put the fact-relative theory in the same boat as the evidence-relative theories with respect to the problems under consideration. The second strategy is to account for proper aggrievement not by appeal to the violation of claims, but rather by appeal to the manifestation of insufficient concern either for an individual’s interests or for their claims. I will consider the latter views in the next section.

III Quality of Will theory

We may turn to the Quality of Will theory — first as it pertains to *Danger Zone* and then to the morality of harm-doing generally.

A. Duty of Care (*Danger Zone*)

The Quality of Will theory of the duty of care takes off from the suggestion, which I shall grant *arguendo*, that a negligent act, insofar as it is really negligent, manifests insufficient concern for

claim. More generally, many would agree with Nagel (1979, p. 31) that “strict liability may have its legal uses but seems irrational as a moral position”; and to the extent that the moral case for strict liability has been made, it is usually by appeal to notions such as agent responsibility (see, e.g., Honoré, 1988) rather than the harmed individual’s status as wronged party — but see Gardner (2001, p. 113).

human interests – that this is what we mean when we call an act “unreasonable”.³⁵ The theory then goes on to identify the particular individuals who are properly aggrieved on account of a negligent act with the particular individuals whose interests have been taken too lightly by the negligent actor. On one version of this theory, the negligent actor violates the claims of all and only those individuals whose interests he took too lightly when he performed the negligent act. On another version, those same individuals constitute the set of individuals properly aggrieved by the negligent act, irrespective of whether they also constitute the set of individuals whose claims were violated. Since nothing turns on it for our purposes, I will simply assume the former version. I will shortly consider whether this theory can be of any use to proponents either of *Danger Zone* or of *Andrews’ Rule*. First, however, I need to say a little bit more about this conception of the duty of care.

Adopting Strawson’s terminology (though modifying his view in certain significant respects), we may recast each claim as a “demand” or “expectation” that the claim-holder’s interests be given their appropriate weight in the practical deliberations of others. The demand is not for a bare emotional or mental state (e.g., the bare state of regard or concern) that may be evidenced by an individual’s conduct. Rather, the concern that we demand of others is distinctly practical and tied closely to the exercise of their agency. It is a demand that others take the fact

³⁵ Some think that tort law’s “objective standard,” which holds an unwitting agent responsible for risks that the similarly situated “reasonable person” would have discerned, precludes the Quality of Will interpretation of tort law’s duty of care. See, e.g., Holmes (1881, Lecture III); Goldberg (2015, pp. 487–90).

that an act may bear on our interests as a reason, of appropriate weight, that counts for or against performance of the act. Moreover, the demand corresponds to our higher-order interest in how our interests are taken into account by others when they are deciding how to act. This higher-order interest is non-instrumental: we have reason to care about how our interests are taken into account in the practical deliberations of others, and not merely because this will tend to affect how well we are treated by others. Additionally, we must *not* say that the demand is that people be *disposed* to take our interests seriously in their deliberations. If a generally considerate loved one acts selfishly on a given occasion, this will count as an objectionable lapse of “regard” or “good will” in the relevant sense, even if the act marked a departure from their settled disposition and can be attributed to exhaustion or stress. Likewise, if my benefactor reliably promotes my interests, but only because she stands to gain, she has not satisfied my demand; for the demand is that others take an act’s effects on my life as itself a reason of appropriate weight bearing on their decision, quite apart from its effects on theirs. Of course, so long as such an individual acts as they would have acted if they had assigned due weight to my interests, they will not be liable for my ensuing losses, since their breach of the duty of care, however real, could not be credited as the cause of those losses.

With this conception of the duty of care in hand, let us see whether it licenses the assignment of rights to all and only members of the danger zone. Recall that, in our opening hypothetical, I may have had every reason to believe that Bertha will be on the other side of town tomorrow, far away from the danger zone (the office building, noon tomorrow). In such a case, can anyone say that my act manifested a failure to take Bertha’s interest seriously? If there is any doubt, let us stipulate that my negligent act was performed precisely in order to help Bertha — that is, I risked harming unknown others so as to benefit Bertha, to whom I am unflinchingly

devoted. In this case, it is not just that my general concern for and devotion to Bertha cannot be called into question, but that the negligent act cannot be thought to constitute a lapse or departure from such settled dispositions. And yet *Danger Zone* holds that I am nonetheless liable to Bertha in such a case, a decision that cannot be justified by the Quality of Will conception.

One might object to this line of reasoning by insisting that my failure to give appropriate weight to the interests of “the building’s occupants, tomorrow” *was* a failure to take Bertha’s interests sufficiently seriously, given that she in fact satisfied the indefinite description. That is, even though I did not fail to take Bertha’s interests seriously under the description “Bertha’s interests,” and even though I could not have known, at the time of action, that Bertha instantiated the danger-making property, all that is needed to propel Bertha into the class of the Properly Aggrieved is a failure to take her interests sufficiently seriously under some description that she in fact satisfies – including the description that she satisfies in virtue of occupying the danger zone.

Once again, the Times Square hypothetical shows that, even if this proposal were correct, it would not help reconcile the Quality of Will account and *Danger Zone*. If my act creates identical risks of harming whoever instantiates the danger-making property and whoever occupies the apartment unit, then if it manifests insufficient concern for one, it must also manifest insufficient concern for the other. In any case, the proposal is not correct. That is, it is not true that all that is needed to propel an individual into the class of the Properly Aggrieved is a failure to take the individual’s interests sufficiently seriously under some description that the individual in fact satisfies.

It is tempting to demonstrate the inadequacy of the proposal by appealing to the notion of *de re* belief: Due to the descriptive character of my belief about the danger zone occupants, I did

not believe *of* Bertha that my act would unreasonably risk harming her. (This assumes, in line with the prevailing orthodoxy, that the belief that I would risk harming the occupants of the danger zone (*whoever they are*) is not a belief about Bertha in particular.) And if I did not believe this *of* Bertha, then I did not manifest a failure to take her interests seriously. However tempting, such a reply would be misguided, or at least incomplete. Suppose that, at the time of the negligent act, I enjoyed perceptual access to the danger zone occupants, including Bertha. However, perched at a distance, Bertha appeared to me as a measly unrecognizable speck at the edge of my visual field. Even if such meager perceptual access entails that I believed of Bertha that my act placed her in harm's way, it would be artificial to draw a distinction between this case, where my access to Bertha was perceptual, and the original case, where my access was exhausted by indefinite description. Whether I "saw her not" or "marked not what I saw," Bertha cannot claim that my act manifested insufficient concern for her interests in particular — her interests, as opposed to the interests of particular others or of all and sundry.³⁶ To be sure, this raises the interesting question of just how somebody must figure in my deliberations in order to make me liable to the charge that I slighted or disrespected them by treating their interests too lightly.³⁷ Rather than provide such a theory, I have supplied a datum that any adequate theory

³⁶ "She saw him not, or marked not if she saw//One among many, though his face was bare."
(Tennyson, *Idylls of the King*).

³⁷ For all that I have said so far, it is possible that *de re* belief is necessary (though not sufficient) for a manifestation of insufficient concern for the interests of a particular person. Cf. Setiya (2023). I am skeptical of such a view, for reasons that will become clear shortly in the discussion of *Andrews' Rule*.

must capture: when Bertha was but an unrecognizable speck on my visual field, then my negligent act did not set back her higher-order interest that I assign appropriate weight to her interests.³⁸ I conclude that, according to *Danger Zone*, a failure to manifest sufficient concern for the interests of an individual is not necessary for violating a duty of care owed to that individual.

Nor is such a failure sufficient for violating the duty of care imposed by *Danger Zone*. If a negligent actor had every reason to believe both that Hector would be in the danger zone and that anyone in the danger zone would be unreasonably endangered, then if the act manifested insufficient concern for anybody's interest, it manifested insufficient concern for Hector's interest. And yet if Hector did not wind up in the danger zone, *Danger Zone* plainly implies that he has not been wronged.

B. Morality (intentional harm-doing)

As with the Interest Theory, the argument concerning the Quality of Will theory generalizes beyond the domain of negligence and beyond the legal context. Let us vary the opening hypothetical, such that I act with the *knowledge* that each of the office building's occupants will suffer significant burns as a result of my action. Like before, we may suppose not only that I thought that Bertha was out of harm's way, but that my devotion to Bertha was unflagging. Nevertheless, if Bertha is in the building and suffers severe harm, then clearly I have wronged her, even though I did not manifest insufficient concern for her interests. Moreover, I

³⁸ More exactly, my negligent act did not set back her higher-order interest if it did not do the same for of a much larger set of individuals, extending well beyond the set of danger zone occupants. I will elaborate on the latter possibility below, though it will obviously be of no help to proponents of *Danger Zone*.

take it that Bertha would be properly aggrieved: as everybody knows, “I did not know it was you” is not an excuse — or, rather, it is not an excuse when, so far as I knew, the occupants of the building were innocents who had done nothing to forfeit their rights. Of course, I do not mean to deny that the sincere exclamation of “I didn’t know it was you,” in the contemplated circumstances, can carry great interpersonal significance. Indeed, it is not hard to imagine that it might save my relationship with Bertha. The lesson we should draw from this, however, is that we should not appeal to the kind of concern that we expect from those with whom we stand in robust personal relationships in order to explain our demands concerning more impersonal modes of mistreatment, such as killing, abducting, defrauding, wounding, falsely imprisoning, thieving, or carelessly endangering. The concern I expect from the far-flung lighthouse keeper (and from everybody else) is not only less than that which I expect from the people in my life, but of an entirely different, impersonal character.

C. Duty of Care (*Andrews’ Rule*)

Let us return to the duty of care. We have considered cases where the negligent actor had reason to believe of the actual occupants of the danger zone that they would not be put in harm’s way. Does it muddy the waters that, in many ordinary cases, the relation between the negligent actor and the actual danger zone occupants are anonymous in character? In fact, reflection on the more common cases only reinforces the poor fit between *Danger Zone* and the focus on the quality of will. If the sole danger zone occupant was not on my radar, such that I could not form any beliefs about her, then it becomes even less plausible to say that my act manifested a failure to take her interests seriously — or, more exactly, it becomes less plausible to say that my act manifested a failure to take her interests seriously in a way that distinguishes her from everybody else.

This last suggestion, that my act manifested an insufficient concern for the teeming multitude, is worth exploring in connection with *Andrews' Rule*. In my example, Bertha ended up in the danger zone even though I had every reason to believe that she would be elsewhere, out of harm's way. Since my act was negligent, it follows that I believed or should have believed that it unreasonably risked harming someone or other. Although it cannot be said that my act displayed a failure to take Bertha's interests seriously, it does manifest something about my attitudes toward the interests of others. At the very least, it manifested a failure to give proper deliberative weight to the mere fact that my act risked harming someone or other — that is, risked harming some human being. The question is whether such a manifestation can render me liable to the resentment of particular others — and if so, which others? In this connection, there are two proposals worth exploring. (As it is not my aim to defend *Andrews' Rule*, I offer these proposals in a tentative spirit.)

For everybody, including those to whom I am unflinchingly devoted, it remains true that, in discounting the interests of my fellow human beings generally, I fail to recognize all the grounds in virtue of which their interests matter. In acting negligently, I display that my concern for them, however great, is contingent on whatever relationship or appraisal happens to ground it. On this suggestion, the attitude I manifested when acting negligently gives everyone occasion for (personal) resentment, and not merely (vicarious) indignation.

But there is a second explanation that might be available to a proponent of *Andrews' Rule*. With respect to most humans, for largely practical reasons, the only basis that I have for taking their interests seriously resides in their humanity. It is not just that I neither stand in special relationships with them nor have any special fondness for them, but that there is nothing about them, over and above their humanity, that might lead me to assign weight to their interests

in my practical deliberations. They are, in short, *merely human* to me. I do not mean to imply that the mere absence of personal acquaintance — social or cognitive — is either necessary or sufficient for membership in this literally undistinguished class, the class of the merely human. For example, there may well be members of my tribe (the tribe of philosophers, say) with whom I am unacquainted who would not qualify as merely human in virtue of the significance I attach to tribe membership. Accordingly, the second proposal is that when an actor fails to give appropriate weight to human interests “as such,” she thereby manifests a failure to assign an appropriate weight to the interests of each and every person who is, to her, merely human. If you are nothing but human to me, then I cannot manifest indifference to human interests without also manifesting indifference to your interests. Of course, my attitudes towards humans may change over time, and by the time that you and I cross paths your species may have won me over. Or maybe you will win me over, lifting yourself out of the class of the merely human by dint of a newfound friendship. Alternatively, my negligent act may have been “out of character” — that is, a fleeting departure from a more settled philanthropic disposition, owing perhaps to fear, shame, anxiety, stress, or outsized concern for another. Either way, when I performed the negligent act I displayed the weight that I then attached to your interests in my practical deliberations. Perhaps this is what Andrews meant when he wrote that a negligent act “is a wrong not only to those who happen to be within the radius of danger but to all who might have been there” — *a wrong, in other words, to all those who might as well have been there, as far as the negligent actor was concerned.*

This line of thought must be qualified in one respect. Somebody may carelessly endanger members of a certain population (e.g., the destitute) only because he thinks, falsely, that members of that population are subhuman, in the sense that they do not warrant the level of

concern that is owing to other human beings. In such a case, we cannot infer from the fact that such an agent has manifested insufficient concern for human interests (as he has) that he manifested insufficient concern for everyone who is merely human to him. Of course, since it would be perverse to let a negligent actor off the hook as a result of their prejudice, this qualification need not be reflected in Andrews' expansive liability rule.³⁹

D. Quality of Will-Interest Theory (hybrid)

A variation on the Quality of Will theory begins with the observation that there is a conceptual relation between violating somebody's claims and being liable to their resentment. Absent an excuse, if you violate somebody's claim, this makes it fitting for them to resent you. Wallace and others (e.g., Rosen 2014, 2015) have sought to explain this connection by marrying the interest theory of claim-rights with the following theory of proper aggrievement ("moral injury"): An act renders an agent liable to the resentment of another only if the act manifests a disregard for the latter's claim-rights. Whereas the previous version of the Quality of Will theory invoked the notion of disregard for the *interests* of others, this version appeals to disregard for the *claim-rights* of others. (For Wallace, given the grounds of claim-rights, to disregard someone's claims is a way of disregarding their interests.) Although I have already criticized Wallace's account of claim-rights, I will set aside those more fundamental objections and register a further difficulty for his theory of proper aggrievement.

For Wallace, the notions of *infringing a claim* and *disregarding a claim* are importantly different. Whereas the infringement of a claim involves impermissible conduct, to disregard

³⁹ Similarly, there may be good institutional reasons not to draw an exception for when the injured party is (e.g.,) the friend of the negligent actor.

somebody's claim-right is to treat it with "indifference" or "contempt," that is, to fail to give the claim its proper weight in one's deliberations (Wallace, 2019, p. 11). Not every infringement of a claim amounts to disregard in this sense; indeed, the very point of the traditional excuses, in Wallace's neo-Strawsonian view, is to show how an impermissible act might not have been the product of such disregard on the part of an agent. Conversely, a claim can be disregarded even when satisfied: "Thus, suppose A [who had promised B not to show up at a certain reception] goes to the gallery, believing that that is where the reception will be happening, and not caring about the fact that she promised B not to attend, but it turns out that the reception was all along scheduled to take place in a different venue on the other side of town" (*ibid.*). While Wallace leaves it open whether B is, "strictly speaking," morally injured in such a case, he implies that his account is consistent with an affirmative answer to the question, and I shall hold him to this position (*ibid.*). Indeed, insofar as resentment is, on Wallace's neo-Strawsonian account, a fitting response to the manifestation of a certain *attitude* of disrespect, insisting on impermissible conduct as a condition of fitting resentment would be ad hoc.⁴⁰

The problem with this theory of proper aggrievement is that in many perfectly ordinary occasions of wrongdoing, the unexcused violation of the claim-rights of any given individual manifests disregard for the rights of that individual only if it also manifests disregard for the rights of many more individuals, individuals whose rights were not violated on the occasion. For example, in a neighborhood of cookie-cutter houses, a thief from out of town may select a certain

⁴⁰ While Rosen (2015, p. 76) posits wrongful conduct as a necessary condition of proper aggrievement, he interprets the condition so weakly that a failure to be appropriately motivated by the bearing of one's act on the interests of another would satisfy the condition.

house to burglarize solely because its lights are turned off or in virtue of its proximity to the highway. In such a case, there is plainly no basis for saying that the wrongdoer's violation of the homeowner's property rights did not manifest an equal disregard for the property rights of a much larger set. With respect to such ordinary instances of wrongdoing, the prevailing practice nevertheless places the burglarized home owner in the privileged position of victim. Yet Wallace's Quality of Will account — like any Quality of Will account — does not (for better or worse) have the resources to restrict such badges of victimhood to the "victim" (i.e., the burglarized homeowner) alone. Although the house singled out for the burglary belonged to a particular individual, *as far as the burglar was concerned, it might as well have belonged to anybody.*

Let us take stock. Our basic question has been to determine when an individual counts as properly aggrieved as a result of another's conduct. One family of views answers this question by appeal to the relation between the individual's interests and the conduct in question. According to these views, if the conduct violates (and/or manifests indifference towards) an obligation that relates in a certain way to the individual's interests — a relation that confers on the individual a claim-right corresponding to the obligation — then the individual belongs to the class of the properly aggrieved. Some views in this family are evidence-relative in that an individual's status as claim-holder is a function of the expected outcome of either the required behavior or the obligation itself. Such views in turn divide between the ex-ante and the ex-post. The ex-ante views hold that a given individual's status as right-holder is a function of the act's expected effects on the interests of that individual. We have seen that such views have unattractive implications concerning who qualifies as properly aggrieved (see *Grand Central*

Station). By contrast, while the ex-post views avoid these implications, they violate basic theoretical constraints — “parity constraints” — governing an obligation’s grounds. Finally, to the extent that interest theorists (or others) posit the agent’s “indifference” to an individual’s claim-rights as a sufficient condition of the latter’s proper aggrievement, they are vulnerable to the charge that, often enough, one cannot manifest indifference to one person’s claims without manifesting indifference to the claims of a much larger set that extends beyond the set of the properly aggrieved, on prevailing understandings of the latter. And in this respect, we have also seen that what is true of indifference to the claim-rights of others is also true of indifference to the interests of others.

Where does this leave us? Let us leave behind Andrews’ expansive conception — not only because of its extreme revisionism, but also from a resistance to the idea that the concern we demand from everyone differs only in degree from that which we expect from our friends, colleagues, and neighbors. I indicated at the start that my favored approach to proper aggrievement is a conventionalist one, on which the selection of the class of victims (with all the attendant privileges, etc.) is at bottom a question of social and institutional policy. The conventionalist theory of proper aggrievement has the same structure as a widely-held theory of punishment.⁴¹ According to the latter theory, the appropriateness of a quantum of punishment in any given case cannot simply be “read off” a characterization of the culpable conduct, including its (actual or expected) harmful consequences; rather, punishment is justified only insofar as it is meted out pursuant to a social or legal policy that is itself justified, i.e., that serves legitimate social goals (e.g., deterring harmful conduct) and fairly distributes benefits and burdens

⁴¹ Scanlon (1998, pp. 262–67) remains an important statement.

(including the burden of being punished). According to the view I am recommending, what is true of *just deserts* is also true of *proper aggrievement*. That is, while an act's actual or expected effects on a certain individual's interests may be sufficient to explain *why* an act is impermissible, we have seen that such effects do not explain who qualifies as a victim of the act. Of course, just as our punishment policies make certain features of wrongful conduct relevant to the determination of appropriate punishment, so too our accountability practices may render an act's effects on a certain individual's interests relevant to whether that individual is properly aggrieved (e.g., by imposing a harm condition on proper aggrievement).

A central aim of this paper has been to prize apart the determinants of right action, on the one hand, and proper aggrievement, on the other. It would be consistent with this aim to recognize a valid set of principles governing proper aggrievement that are "*non-conventional*" in the sense of not depending for their validity on the imprimatur of social or institutional norms or practices. It would also be consistent to adopt a skeptical view, one that condemns social and institutional practices that confer upon select individuals the status of the properly aggrieved. Arguing against such skepticism would require a more detailed inquiry concerning the interests and values bearing on accountability practices than I can here undertake. I will instead close by gesturing towards reasons for thinking that any successful justification of proper aggrievement must be conventional in character.

Variation in conceptions of proper aggrievement across communities is well attested. Consider the following teaching of Maimonides concerning forgiveness:

Even if a wrongdoer has compensated his victim, he must still appease him and seek his forgiveness. Even if a person only upset his fellow by saying offensive things, he must still appease him and seek forgiveness... If the victim is not appeased, the wrongdoer

should repeat the process and ask forgiveness a second and third time. If the victim still does not want to grant forgiveness, the wrongdoer may leave him be and need not pursue the matter any further. On the contrary, the person who refuses to grant forgiveness is alone deemed the offending party. (*Laws of Repentance* (2:9))

On this view, still prevalent in many Jewish communities, the protracted withholding of sought-after forgiveness not only extinguishes the offender's reasons to apologize but also deprives the victim of many of the powers and privileges of victimhood, including the power to forgive. This view is not universally held, and in our own secular, liberal societies many would grant a victim of a sufficiently serious wrong the moral prerogative to withhold even sought-after forgiveness in perpetuity, much less the moral power to forgive after a protracted period of unwillingness. Still, the Maimonidean approach marks a legitimate way of striking a balance between important values — the recognition of the gravity of an interpersonal wrong, on the one hand, and of the offender's own inalienable dignity, on the other — a dignity that should preclude us from viewing his past choices as forever determinative of his character or his present ones as of limited significance. Of course, the conventionalist holds that the variation in question is not merely sociological but also normative: individuals have normative reason to conform to the rules of their own practice, and not some other, provided that their own is both sufficiently entrenched and legitimate (“good enough”).⁴² I offer the above example as an intuitively plausible candidate

⁴² Of course, the conventionalist need not deny that some legitimate practices are better than others (i.e., can distinguish between *legitimate* practices and *optimal* ones), and may also hold that one can have reason to conform to a not-yet-existing practice, provided that one has sufficient reason for thinking that one's conformity may help bring it about.

for just such a variable feature of an accountability practice. To put it very crudely, it will seem true to anyone familiar with both communities “from the inside” that the unforgiving victim in Borough Park, NY can lose his moral power to forgive in ways that his counterpart in Greenwich, CT cannot.

The motivation for conventionalism about accountability is not limited to such intuitions, and can be given a theoretical basis. To see this, consider the naturalness of the connection drawn by Maimonides between forgiveness and appeasement. Without arguing the case, I submit that we cannot intelligibly conceive of a practice of apology and forgiveness that is not a component of a wider practice of blame and resentment. And I will also suggest that the appropriateness of episodes of blame or resentment depends on the salutary consequences of a more general practice licensing such episodes.

Although I will not offer an account of blame, I must clarify that the kind I have in mind — the kind of blame that might be deemed appropriate for victims alone — is both hot (angry or resentful) and overt (confrontational or otherwise expressed). That people are averse to becoming the object of such blame — that being such an object is closely tied to fear, shame, and other undesirable forms of diminishment and marginalization — is perhaps the least controversial thing one can say about it.⁴³ However, this is all we need to say in order to sustain an analogy with punishment: if it is the knowing infliction of suffering that creates a special justificatory burden for the practice of punishment, then episodes of blame should be charged with that burden as well. To be sure, there may be important differences between punishment

⁴³ See McKenna (2012, pp. 134–41); Pereboom (2014, p. 180); Holmgren (2014, p. 169); but see Shoemaker (2018).

and blame with respect to the suffering they produce. For example, while the paradigmatic forms of punishment involve treatment (e.g., lashes or confinement) that would be undesirable independently of the context of punishment, what we dread about becoming the object of blame is difficult to specify without referring to the blame itself. However, I do not see why the justificatory burden that applies to punishment should be sensitive to this difference, nor to differences in the characteristic aims of blamers and punishers in their respective inflictions of suffering. While it is important to distinguish between punishment and blame, it is equally important to recognize that both are forms of “harsh treatment”.⁴⁴ Moreover, insofar as it involves the manifestation of a desire to exact “payback” in response to a perceived wrong (a venerable, if controversial, view), this is yet further reason to hold blame to the same justificatory standards as the exaction itself.

Like many non-retributivists, I hold that punishment can be independently justified, if at all, only by values realized at the “system level,” not the individual case. It is (knowledge of) a general policy or practice to punish that serves to deter wrongful conduct, for example; carrying the policy out in any given case is (morally speaking) a cost of adopting the policy, not its point, and is justified only insofar as it is called for by the justified policy. Moreover, in many contexts a standing personal policy that can be attributed to the punisher alone, rather than to members of a broader group, would be insufficiently deterring to justify the infliction of suffering. For similar reasons, the remonstrations, and angry feelings, that are part-and-parcel of garden-variety blame can be justified, if at all, only as instances of more general norm-governed patterns of

⁴⁴ Note that I am not assuming that we have reason to want to avoid every phenomenologically unpleasant mental state (e.g., remorse). See Achs (forthcoming) for a defense of “guilt-tripping”.

behavior that serve socially useful functions for a group.⁴⁵ Of course, I do not deny that individual episodes of blame may both gratify the victim and transform the wrongdoer for the better; however, the same is no less true of punishment.

Considerations of fairness also favor conditioning justifiable blame on conformity to a blaming practice. In the case of punishment, many non-retributivists hold that even if a punishing practice effectively deters wrongful conduct, it would be unfair (and hence unacceptable) if it does not give people “adequate opportunity” to avoid the punishment.⁴⁶ Suppose that a society adopts a policy of punishing those who engage in a certain act (ψ), that the policy has retroactive application — that is, calls for punishing people who are found to have ψ 'd prior to the policy's adoption — and that the policy can be rationalized on deterrence grounds. The mere fact that an agent who ψ 'd prior to the policy's adoption could have refrained from doing so at minimal cost to herself by no means implies that the agent had “adequate opportunity” to avoid the punishment, in the relevant sense. Indeed, on a commonly held non-retributivist view, such retroactive punishment would be unfair even if the agent not only could have refrained from ψ -

⁴⁵ On the deterrent function of blame, compare Fricker (2016, p. 174): “[T]he prospect of being on the receiving end of Communicative Blame is part of what keeps our behaviour in check...Communicative Blame in one or another form is an entirely everyday mechanism by which we hold each other accountable. We don't like being blamed, being found fault with, and especially not if that means being on the receiving end of some emotional flak, whether of the heated or chilly sort.”

⁴⁶ See Scanlon (1998) and Kolodny (2023, pp. 49–55).

ing at the earlier time, but was morally obligated to do so.⁴⁷ On this view, fairness requires not only that punishment be meted out only pursuant to a more general practice, but also that agents be spared if they were not adequately informed of the practice at the time of action. As before, if this holds for punishment, it should hold for blame as well. While I have not yet made the case for conventionalism about proper aggrievement, I hope I have said enough both to indicate how such an argument would go and to demonstrate its appeal.

Works Cited

Achs, R. (forth.) In defense of guilt-tripping. *Philosophy & Phenomenological*

Research. <https://doi.org/10.1111/phpr.13009>

Coleman, J. L. (2008). Doing away with tort law. *Loyola Law Review*, *41*, 1149–1170.

<https://digitalcommons.lmu.edu/llr/vol41/iss4/2>

Cornell, N. (2015). Wrongs, rights, and third parties. *Philosophy & Public Affairs*, *43*, 109–143.

<https://doi.org/10.1111/papa.12054>

⁴⁷ This is not controverted by Kolodny’s “Wrongful Benefit Principle” (2023, p. 52) which holds that “the fact that someone had a duty to X *can* itself *contribute* to making it the case that their opportunity to avoid force by X-ing was adequate” (emphasis added). As his discussion reveals, this principle speaks only to the *costs* of refraining from whatever act rendered someone liable to force; however, the possibility of retroactive punishment makes it clear that the relevant notion of “adequate opportunity” takes into account more than avoidance costs. I also note that while the “adequate opportunity to avoid” requirement considered by Kolodny applies only to bodily invasions, I have suggested that restricting it to such invasions amounts to the fetishization of the body.

- Cowan, T. (1938). The riddle of the *Palsgraf* case. *Minnesota Law Review*, 23, 46.
<https://scholarship.law.umn.edu/mlr/2424>
- Darwall, S. (2006). *The second-person standpoint*. Harvard University Press.
- Frick, J. (2015). Contractualism and social risk. *Philosophy & Public Affairs*, 43, 175–223.
<https://doi.org/10.1111/papa.12058>
- Fricke, M. (2016). What’s the point of blame? *Nous*, 50, 165–183.
<https://www.jstor.org/stable/43828961>
- Gardner, J. (2001). Obligations and outcomes in the law of torts. In P. Kane (Ed.), *Relating to responsibility*. Hart Publishing. <https://ssrn.com/abstract=1397133>
- Goldberg, J., & Zipursky, B. (2021). *Recognizing wrongs*. Harvard University Press.
- Goldberg, J. (2015). Inexcusable wrongs. *California Law Review*, 103, 467–512.
<https://www.jstor.org/stable/24758482>
- Hart, H. L. A. (1982). Legal rights. In *Essays on Bentham* (pp. 162–193). Clarendon Press.
- Hirsch, E. (forthcoming). Negligence and foreseeable damage: The Talmudic eye of vigilance. In P. Franks & Y. Melamed, (Eds.), *Oxford handbook of Jewish philosophy*. Oxford University Press.
- Holmes, O. W. (1881/2009). *The common law*. Harvard University Press.
- Holmgren, M. (2014). A moral assessment of Strawson’s retributive reactive attitudes. In D. Shoemaker & N. Tognazzini (Eds.), *Oxford studies in agency and responsibility*, Vol. 2, pp. 165–186. Oxford University Press.
- Honoré, T. (1988). Responsibility and luck. *Law Quarterly Review*, 104, 530–553.
- Horton, J. (2017). Aggregation, complaints, and risk. *Philosophy & Public Affairs*, 45, 54–81.
<https://doi.org/10.1111/papa.12084>

- Jonker, J. (2019). Contractualist justification and the question of direction. *Legal Theory*, 25, 200–224. <https://doi.org/10.1017/S1352325219000089>
- Kolodny, N. (2023). *The pecking order*. Harvard University Press.
- McKenna, M. (2012). *Conversation and responsibility*. Oxford University Press.
- Murphy, L. (2021). Nonlegislative justification. In J. McMahan, T. Campbell, J. Goodrich, & K. Ramakrishnan (Eds.), *Principles and persons*, pp. 247–274. Oxford University Press. <https://doi.org/10.1093/oso/9780192893994.003.0012>
- Nagel, T. (1979). Moral luck. In *Mortal questions*, pp. 24–38. Cambridge University Press. <https://doi.org/10.1017/CBO9781107341050.005>
- Pereboom, D. (2014). *Free will, agency, and meaning in life*. Oxford University Press.
- Perry, S. (2014). Torts, rights, and risks. In J. Oberdiek (Ed.), *Philosophical foundations of tort law*, pp. 38–64. Oxford University Press.
- Prosser, W. (1953). *Palsgraf revisited*. *Michigan Law Review*, 52. <https://repository.law.umich.edu/mlr/vol52/iss1/2>
- Ramakrishnan, K. (m.s.). The problem with *Palsgraf*.
- Raz, J. (1984). Legal rights. *Oxford Journal of Legal Studies*, 4(1), 1–21. <https://doi.org/10.1093/ojls/4.1.1>
- (1988). *Morality of freedom*. Oxford University Press.
- Ripstein, A. (2016). *Private wrongs*. Harvard University Press.
- Rosen, G. (2014). Culpability and duress. *Aristotelian Society*, supplementary volume.
- (2015). The Alethic conception of moral responsibility. In R. Clarke, M. McKenna, & A. M. Smith (Eds.), *The nature of moral responsibility: New essays*, pp. 65–88. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199998074.001.0001>

- Scanlon, T. M. (1998). *What we owe to each other*. Belknap Press.
- Schwartz, G. (1999). Cardozo as tort lawmaker. *DePaul Law Review*, 49, 305–318.
<https://via.library.depaul.edu/law-review/vol49/iss2/5>
- Seavey, W. (1927). Negligence—Subjective or objective? *Harvard Law Review*, 41, 1–28.
<https://www.jstor.org/stable/1330404>
- Setiya, K. (2023). Other people. In S. Buss & N. Theunissen (Eds.), *Rethinking the value of humanity*, pp. 314–336. Oxford University Press.
<https://doi.org/10.1093/oso/9780197539361.003.0014>
- Shoemaker, D. (2018). You oughta know! Defending angry blame. In M. Cherry & O. Flanagan (Eds.), *The moral psychology of anger*. Rowman & Littlefield.
- Smith, H. (2011). Modularity and morality in the law of torts. *Journal of Tort Law*, 4.
<http://nrs.harvard.edu/urn-3:HUL.InstRepos:8789603>
- Strawson, P. F. (1962). Freedom and resentment. *Proceedings of the British Academy*, 48, 187–211. <https://www.thebritishacademy.ac.uk/publishing/proceedings-british-academy/48/strawson/>
- Thompson, M. (2004). What is it to wrong someone? A puzzle about justice. In R. J. Wallace, P. Pettit, S. Scheffler, & M. Smith (Eds.), *Reason and value*, pp. 333–384. Oxford University Press. <https://sites.pitt.edu/~mthomps/raz.pdf>
- Thomson, J. J. (1992). *Realm of rights*. Harvard University Press.
- Wallace, R. J. (2019). *The moral nexus*. Princeton University Press.
- Weinrib, E. (2003). *The idea of private law*. Oxford University Press.
- Zipursky, B. (1998). Rights, wrongs, and recourse in the law of torts. *Vanderbilt Law Review*, 51. <https://scholarship.law.vanderbilt.edu/vlr/vol51/iss1/1>

