2012

Aggregation and Law

Eric A. Posner

Ariel Porat

THE YALE LAW JOURNAL

ARIEL PORAT & ERIC A. POSNER

Aggregation and Law

ABSTRACT. If a plaintiff brings two claims, each with a 0.4 probability of being valid, the plaintiff will usually lose, even if the claims are based on independent events, and thus the probability of at least one of the claims being valid is 0.64. If a plaintiff brings two independent claims, and neither of them alleges misconduct sufficient to justify a remedy, the plaintiff will usually lose, even if the claims jointly allege sufficient wrongdoing to justify a remedy. Thus, as a general rule, courts refuse to engage in what we call factual aggregation (the first case) and normative aggregation (the second case), as well as other forms of aggregation that we identify. Yet we show numerous exceptions to this rule in private and public law. Notably, in public law, the hybrid rights doctrine permits courts to aggregate two weak constitutional claims as long as one involves free exercise of religion. In private law, certain tort and contract doctrines also permit aggregation. We criticize the courts' inconsistent approaches to aggregation, and propose conditions under which courts should (and should not) aggregate.

ARTICLE CONTENTS

INTRODUCTION

Suppose you are invited to dinner by a friend. You are a bit tired, but not extremely tired, so that reason by itself would not make you decline the invitation. You also want to spend the evening with your family, but this reason standing alone would not convince you to stay at home. Finally, you are also a bit pressed for time because you need to prepare a lecture for tomorrow, but once again you would not miss the dinner for that reason only. It is quite plausible that even if none of these reasons, standing alone, is sufficient for you to decline the invitation, the aggregation of all three reasons would be sufficient. Nevertheless, we suspect that most people, while aggregating the three reasons for themselves and declining the invitation, would not say to their friend that they cannot come to dinner because: (1) they are tired; (2) they want to spend the evening with their family; and (3) they need to prepare a lecture for tomorrow and are therefore pressed for time. They would instead choose the strongest of the three reasons and provide it as the sole reason for declining the invitation.

Consider another possibility. Your friend invites you to dinner a week in advance. Peering into the future, you predict that with some (low) probability you will be too tired, that with some (low) probability your children will need help with their homework, and that with some (low) probability you will need to prepare for work on the following day. You realize that while each event will individually come to pass with low probability, the probability that at least one of the events will come to pass is quite high. Even so, you would not say to your friend (if you want to keep your friendship) that while each reason you have for turning down the invitation is low probability, they are jointly high probability. Most likely, you would turn down the invitation on the basis of the most probable reason.

These puzzles, which we call "aggregation puzzles," have counterparts in the law. Consider a plaintiff who brings two separate claims against the defendant arising from the same event, a car accident. The plaintiff argues that the defendant committed a strict liability tort by driving an inherently dangerous vehicle and, in the alternative, that the defendant caused a tort through negligent driving. To win on the strict liability claim, the plaintiff must prove that the vehicle was inherently dangerous, but the plaintiff can provide evidence to show only a 40% probability of inherent dangerousness. In addition, the plaintiff can show only a 40% probability of negligence. A court would hold against the plaintiff on both claims because she cannot meet the 50% preponderance-of-the-evidence threshold for either claim. However, if the

4

two claims are independent,[1] the plaintiff can show a 64% probability that the defendant committed either one tort or the other.[2] Yet courts do not permit this type of *cross-claim factual aggregation*.

For another example, this one involving claims arising from two distinct events, consider a plaintiff who can prove with 40% probability that the defendant engaged in a material breach of a contract, and also can prove with 40% probability that the defendant engaged in fraudulent misrepresentation to induce the creation of the contract. Under either theory, the plaintiff would be entitled to rescission of the contract. Yet again, although the probability that at least one claim is valid is 64% (assuming that the claims are independent), the plaintiff would lose, because courts do not permit cross-claim factual aggregation.

A second type of factual aggregation, which we will call *cross-element factual aggregation*, occurs across the elements that make up a single claim. For example, a negligence claim has elements of fault and causation, each of which the plaintiff must prove to establish liability. Say the plaintiff can prove each with 60% probability. If the probabilities of the two elements are not aggregated, the plaintiff wins, because the probability of each element being valid is higher than 50%, and so the plaintiff has established all requisite elements of her claim. But if the probabilities of the elements are aggregated, the plaintiff loses, because the probability that the defendant was at fault *and* caused the harm is only 36%.[3] There is uncertainty about whether and to what extent courts engage in cross-element factual aggregation.

A third type of factual aggregation, which we call *within-element factual aggregation*, occurs when the plaintiff proposes several alternative factual theories that satisfy a single element of a claim. Say, for example, that the plaintiff suing for negligence has two distinct theories for why the defendant was at fault, and she can prove that each theory is true with 40% probability. If the probabilities of the two theories are not aggregated, the plaintiff will fail to establish the element of fault by preponderance of the evidence. If the probabilities are aggregated, however, the plaintiff will successfully establish fault, because the probability that one theory or the other is true is 64%. Unlike cross-claim factual aggregation, within-element factual aggregation is routine in American jurisprudence.

---

1. Two claims are independent if the probability of the two claims being valid is independent; that is, if one claim is valid, it does not affect the probability of the other claim being valid. So long as two claims are not *fully* dependent, aggregation is not precluded. For further discussion of dependence, see *infra* notes 37-39 and accompanying text.

2. The probability of at least one claim being valid is $1 - (1 - 0.4)^2 = 0.64$.

3. $0.6^2 = 0.36$.

Aggregation does not always require uncertainty. Suppose that a plaintiff can show with certainty that the defendant induced the creation of a contract with misleading remarks that fall just short of fraudulent misrepresentation, and that the defendant subsequently engaged in a breach that falls just short of material. A court would typically not grant rescission, but not because the plaintiff failed to establish facts with sufficient probability. The plaintiff would lose because, for each claim, she failed to allege sufficient wrongdoing. Yet one could argue that—even if each is insufficient standing alone—the two bad acts together justify rescission.

We will call this type of aggregation *normative aggregation*. Like factual aggregation, it has cross-claim, cross-element, and within-element variations. Combining the independently insufficient claims for fraudulent misrepresentation and material breach described above, for example, would entail cross-claim normative aggregation. The normative weights of the two claims, neither of which standing alone is sufficient to establish liability, would be aggregated, and the combined weight of the two claims would establish liability.

Although courts do not usually permit this type of aggregation, such aggregation does occur in one important class of cases. When a neutral and generally applicable statute burdens religious exercise alone, it does not violate the First Amendment. But if the law simultaneously burdens another constitutional right, such as the right to free speech—even if that burden is not sufficient to violate that constitutional right independently—the law may nonetheless be overturned because of the aggregate burden it imposes on two constitutional rights.[4]

One can also imagine cases that share aspects of cross-claim factual aggregation and cross-claim normative aggregation. Suppose that the plaintiff can prove material breach with 40% probability, and can prove with certainty that the defendant induced the creation of the contract with misleading remarks that fall just short of fraudulent misrepresentation. One might argue that the plaintiff should be entitled to rescind the contract, but courts do not permit this type of *cross-claim mixed aggregation*.

Another type of aggregation takes place across persons. Suppose that a firm pollutes the air, and ten nearby residents claim that they were injured by the pollution. Each resident can show that she breathed in the pollution and that her medical condition deteriorated after the pollution, but all residents suffer from preexisting respiratory ailments, and thus cannot show with probability above 50% that the pollution, rather than their preexisting conditions, caused

---

4. *See infra* notes 121-128 and accompanying text.

their harm. They would therefore lose their cases. Yet if each resident could show that the probability that the pollution exacerbated her medical condition is, say, 10%, then the residents can collectively prove that the probability that at least one of them was injured was greater than 50%, and therefore that the firm should pay damages[5] (although not necessarily everyone's damages—an issue we will address later). We will call this type of aggregation *cross-person aggregation*. Cross-person aggregation could be factual, as in the preceding example, but could also be normative or mixed.[6]

These examples illustrate an important vulnerability at the heart of the law. They reflect the fact that law relies on legal categories that organize the judicial treatment of disputes. These categories operate at different levels of generality, including *bodies of law* (tort, contract), *claims* (strict liability, negligence), and *elements* (offer, acceptance, breach, harm). These categories are important, and it is hard to imagine how the law would work without them. But they also require courts to disregard information that is relevant to an overall evaluation of the asserted wrongdoing.

This happens in the ways we have illustrated. First, where two claims concern a single act by the defendant, some of the factual information that is relevant for evaluating the defendant's wrongdoing may need to be disregarded when one claim is evaluated, and other factual information disregarded when the other claim is evaluated. An act that is not clearly a strict liability tort and at the same time not clearly a negligence tort may nonetheless clearly be one or the other (assuming that the strict liability and negligence claims are at least partially independent) and thus a wrongful act that should entitle the victim to a remedy. A similar phenomenon transpires when the two (or more) claims relate to two (or more) events, and each event is considered separately, isolated from one another. Even if—for each event standing alone—the plaintiff cannot establish that the defendant caused the harm, the plaintiff may nonetheless be able to establish that the defendant caused the harm in *at least one* of the events. Conversely, even if in each event standing alone the plaintiff can establish that the defendant caused one harm, he may not be able to establish that the

---

5. The probability that at least one resident was injured is $1 - (1 - 0.1)^{10} \approx 0.65$.

6. The phrase "aggregation" could have various meanings in different contexts. In particular, the law allows a type of aggregation for evidentiary purposes through the prior-acts and similar-crimes doctrines applied in criminal law, *see infra* notes 93-102 and accompanying text, according to which past behavior of the accused could serve as evidence to prove his guilt in the present case. What is typical of this type of aggregation is the *dependence* between the accused's different misbehaviors; a defendant is arguably more likely to have committed a crime given his demonstrated involvement in past crimes. Our focus instead is on aggregation of *independent* claims, although we admit that sometimes the distinction between the two types of aggregation is blurred.

defendant caused both harms. In this way, as we explore further below, aggregation can lead to less liability in some cases.

Second, the law relies heavily on thresholds even when wrongdoing is typically a continuous variable.[7] A plaintiff must reach one normative threshold along the continuum of increasingly severe wrongdoing to show fraudulent misrepresentation and another normative threshold to show material breach. But where an event falls just short of the thresholds in two separate legal dimensions, or two events individually fall short of the threshold, the threshold may be exceeded when those dimensions, or events, are aggregated. The defendant who does not quite engage in fraudulent misrepresentation and does not quite engage in material breach may nonetheless have acted wrongfully in her overall treatment of the plaintiff. Conversely, a criminal defendant prosecuted for assault who falls just short of the thresholds for both the insanity defense and a claim of self-defense may not be blameworthy enough to justify conviction.

Third, the law generally treats individuals as the unit of analysis, even though wrongdoing can often be probabilistic, in a sense transcending individuals. The point is not just that a firm that causes a small amount of harm to a large number of people may escape liability because no individual possesses a sufficient incentive to bring suit. This is a familiar problem, one that is addressed by the class action system. The problem is that even if each individual faced zero legal costs, she would lose her case. The probability that she was harmed is low or, alternatively, does not quite reach the normative threshold for each individual, but across many persons, that probability becomes significant.

Each of these cases bears a family resemblance to the others. They all stem from the problem of aggregating two types of things: factual information and normative weight. In the bulk of this Article, we will examine additional examples of how such aggregation might take place in torts, contracts, criminal law, and public law. We will also identify actual instances of aggregation in several of those fields and provide some tentative proposals for reforming how courts incorporate aggregation. Our focus will be general explanations and

---

7. See, for example, John E. Coons, *Approaches to Court Imposed Compromise—The Uses of Doubt and Reason*, 58 Nw. U. L. Rev. 750, 751 (1964), which argues that the law typically requires judges to choose winners and losers on an "all-or-nothing" basis even when the outcomes are harsh. Coons calls for court-imposed compromises based on the idea that there is a continuum of solutions between the two polarized ones: "[I]mposed compromise shall mean the apportionment of right and duty between opposed litigants by a court according to a quantitative standard that is not limited to the favoring of one party to the exclusion of his adversary." *Id.* at 753.

proposals that apply to aggregation in all fields of law. We summarize our conclusions here.

All of the cases reflect a familiar rules/standards tradeoff. The law disaggregates in order to reduce decision costs for courts and other decisionmakers, including ordinary people and firms who want to obey the law. The basic breakdown of wrongdoing into bodies of law, and then those bodies of law into claims, and those claims into elements, brings a regimented clarity to the process of learning and applying the law. But the disaggregation of wrongdoing into a series of rules comes at a cost: morally relevant information is lost.

To some extent, the law already recognizes this problem. Certain doctrines permit courts in certain cases to re-aggregate disaggregated claims. We will discuss examples later, but for now a few such examples that might be cited are the alternative liability rule and market share liability doctrine in tort law, the unconscionability doctrine in contract law, and the hybrid rights doctrine for the Free Exercise Clause in constitutional law. These doctrines permit courts to aggregate claims that would otherwise be kept separate under more conventional types of legal analysis.

However, we will argue that, because of its failure to aggregate more broadly or consistently, the law falls short in many significant respects, some of them illustrated by our examples above. Our minimal goal is to propose "re-aggregation doctrines" that permit courts to aggregate factual and normative claims where doing so does not create excessive confusion. We show that aggregation would often lead to more liability compared to nonaggregation, but occasionally the reverse is true and aggregation would lead to less liability. Our more ambitious goal is to suggest general parameters for the optimal level of aggregation in the law.

The Article proceeds as follows. Parts I through IV analyze the nonaggregation problem in tort law, contract law, criminal law, and public law, respectively. Part V discusses explanations and justifications for courts' refusal to aggregate, offers a theory for analyzing aggregation problems in the law, and proposes methods of implementation. The Conclusion summarizes the discussion.[8]

---

8.   Aggregation has largely been ignored by legal writers. A notable exception is Saul Levmore, *Conjunction and Aggregation*, 99 MICH. L. REV. 723 (2001). Levmore's discussion, however, is limited to factual aggregation and is focused on tort law, specifically on factual aggregation across the elements of the same cause of action. *See infra* discussion accompanying notes 33-35. Some parts of Levmore's analysis—in particular, his discussion of implementation difficulties in aggregation—could be relevant to some types of factual aggregation that we discuss, but not to others. Levmore, *supra*, at 726-33. Another exception is Alon Harel &

I. TORT LAW

A. Factual Aggregation

Factual aggregation in tort law is common for determining whether the defendant committed a particular element of a claim, like fault or causation in a negligence claim. Take, for example, a case where "a car parked at the curb by the defendant begins to roll downhill" and hits the plaintiff, and the reason for the car's roll could be that the defendant "either failed to set the brakes or failed to cut the wheels properly against the curb, or failed to put the car in parking gear." It would be common practice for the court or jury to find that the defendant's fault caused the accident even without knowing what the defendant's faulty behavior was, as long as the aggregated probability that *some* faulty behavior of the defendant caused the harm is sufficiently high.[9] In this way, courts allow what we call *within-element factual aggregation* for proving

---

Ariel Porat, *Aggregating Probabilities Across Cases: Criminal Responsibility for Unspecified Offenses*, 94 MINN. L. REV. 261 (2009). Harel and Porat analyze factual aggregation in criminal law and focus on situations where the accused is charged with two or more separate offenses and none of them can be proven beyond a reasonable doubt, but there is no reasonable doubt that he committed at least one of them. *Id.* at 264-65. Frederick Schauer and Richard Zeckhauser propose aggregating probabilities across cases outside the judicial context. *See* Frederick Schauer & Richard Zeckhauser, *On the Degree of Confidence for Adverse Decisions*, 25 J. LEGAL STUD. 27, 41-51 (1996). They argue that it would make sense for a school to dismiss a teacher against whom several complaints of sexual harassment had been made in the past, even if each complaint, considered independently, would not constitute sufficient reason for dismissal. *Id.* at 44. Schauer and Zeckhauser maintain, however, that such an argument is inapplicable to criminal proceedings. *Id.* at 41, 46, 52. Daryl Levinson considers cases where remedies like school desegregation decrees are based on a kind of aggregation — of multiple past behaviors by government entities (a kind of cross-claim factual aggregation) against multiple persons (a kind of cross-person aggregation). *See* Daryl J. Levinson, *Framing Transactions in Constitutional Law*, 111 YALE L.J. 1311, 1354-55 (2002). Finally, a related literature exists on the problem of multimember decisionmaking bodies. *See, e.g.*, Jonathan Remy Nash, *A Context-Sensitive Voting Protocol Paradigm for Multimember Courts*, 56 STAN. L. REV. 75 (2003) (discussing decisionmaking by multimember judicial panels); Hayden J. Trubitt, *Patchwork Verdicts, Different-Jurors Verdicts, and American Jury Theory: Whether Verdicts Are Invalidated by Juror Disagreement on Issues*, 36 OKLA. L. REV. 473 (1983) (discussing jury verdicts where jurors agree on the outcome but disagree on the facts). Unlike this literature, we assume a single judicial decisionmaker. For a factfinding theory that opposes factual aggregations of any kind, see Kevin M. Clermont, *Death of Paradox: The Killer Logic Beneath the Standards of Proof*, 88 NOTRE DAME L. REV. (forthcoming 2012), http://ssrn.com/abstract=1986346, and in particular footnote 124, which relates directly to our Article and proposes that the paradoxes that motivate our Article do not exist.

9.   For this example and others, see DAN B. DOBBS, THE LAW OF TORTS § 154 (2000).

fault. Courts, however, do not engage in *cross-claim factual aggregation*. Consider the following example:

> *Example I.1. The Inherently Dangerous Vehicle.* The defendant hit the plaintiff while driving. The plaintiff argues that the defendant is strictly liable because he was driving an inherently dangerous vehicle and, in the alternative, that the defendant caused the harm by his negligent driving. The plaintiff, however, cannot establish his claims by the preponderance of the evidence. Specifically, the plaintiff can only show that the probability of each claim is 40%.

Nonaggregation results in both claims being rejected. If the court deciding the case instead aggregates the two claims, and it determines that those claims are independent, it will impose liability on the defendant. The probability that the defendant *wrongfully* hit the plaintiff would be 64%, enough to satisfy the preponderance-of-the-evidence standard.[10] Courts, however, do not aggregate in cases like Example I.1.[11] As a result, defendants escape liability even when the probability that they wrongfully harmed the plaintiff is greater than 50%, simply because the plaintiff cannot establish what *exactly* the wrong committed by the defendant was.

The next two examples represent a cross-claim factual aggregation *relating to two separate events*.

> *Example I.2. Injury in the Hospital: Two Events, One Injury.* The plaintiff was admitted to the hospital while having a heart attack. Initially, the doctor at the emergency room misdiagnosed him and sent him home. When the plaintiff returned two days later, another doctor in the cardiology department allegedly gave him poor treatment. The plaintiff did not fully recover. He sues the hospital for vicarious liability, arguing that the two doctors were negligent, and that each doctor's negligence is a but-for cause of his injury. The evidence before the court indicates

---

10. The probability that either claim is valid is $1 - (1 - 0.4)^2 = 0.64$.

11. In *Candler General Hospital, Inc. v. McNorrill*, 354 S.E.2d 872 (Ga. Ct. App. 1987), the plaintiff allegedly fell and hurt his knee while being transferred by a nurse from a stretcher to a wheelchair. *Id.* at 873. The plaintiff argued that the hospital was vicariously liable for the nurse's negligence, as well as directly liable for negligence due to the inadequacy of the equipment and personnel in the emergency room. *Id.* at 874. The court ordered that both claims should be presented to a jury, but it did not contemplate aggregation of the claims. *Id.* at 876-77. It is quite possible that a jury would hold in favor of the hospital if both claims were examined separately but would impose liability if the claims were aggregated.

that the probability that each of the doctors caused the harm negligently is only 40%.

In contrast to Example I.1, in Example I.2 there are two separate events occurring at different times and places, and the same defendant (the hospital) could be (vicariously) liable for their injurious effects. If the two claims relating to the two events are evaluated separately, liability should not be imposed. If instead the two claims are aggregated, the court should hold the defendant liable, as long as the events are independent, even though it cannot determine which of the two events was the wrongful one.

It seems that courts would not aggregate in cases represented by Example I.2. However, two special doctrines of tort liability arguably could allow aggregation in multiple-event, one-injury cases. First, courts have been willing to impose liability on hospitals and their employees when a plaintiff has established that he suffered harm from the negligence of one of the hospital's employees, even if the identity of the specific employee who negligently caused the harm remained unknown. Courts have imposed liability in these cases under the doctrine of res ipsa loquitur.[12] Although these cases usually involve one event,[13] the doctrine might conceivably apply to two events if a plaintiff could establish that he suffered harm from an unidentified employee's negligence during one of those events.[14]

Second, courts might be willing to impose liability on hospitals and their employees under the "alternative liability rule" from *Summers v. Tice*[15] if a plaintiff could establish that those employees were each negligent, even if the plaintiff could not prove which employee's negligence caused the harm at issue. Under the alternative liability rule, as first codified in the *Restatement (Second)*

---

12. For example, in *Fieux v. Cardiovascular & Thoracic Clinic, P.C.*, 978 P.2d 429 (Or. Ct. App. 1999), a clamp was left behind on the plaintiff's heart during a surgery. *Id.* at 431. The plaintiff could not prove which member of the medical staff, composed of three nurses and one surgeon, was negligent. He relied on res ipsa loquitur to infer negligence. *Id.* The court permitted the plaintiff to proceed to trial against several defendants on that theory, overturning the trial court's directed verdict for the hospital. *Id.*

13. *See, e.g.*, *id.*

14. For example, had the plaintiff in *Fieux* undergone two consecutive but distinct procedures that involved heart clamps, he could have attempted to demonstrate that the harm he suffered must have occurred because of some employee's negligence during one of the two procedures.

15. 199 P.2d 1 (Cal. 1948) (imposing liability on two hunters for the injury one of them caused the plaintiff, when both hunters negligently shot in the plaintiff's direction, and the identity of the one who actually injured the plaintiff was not established).

of Torts,[16] and later in the *Restatement (Third) of Torts: Liability for Physical and Emotional Harm*,[17] when the plaintiff can establish that several defendants acted tortiously toward him, and that at least one of them caused his harm, each of the defendants bears the burden to show that it was not his tortious conduct that caused the harm. Although the alternative liability rule, like res ipsa loquitur, has traditionally been applied to cases of one event during which multiple defendants act simultaneously, its logic also seems to apply to multiple-event cases in which defendants act at different moments in time, as long as the tortious conduct of each defendant can be proven by the preponderance of the evidence. Both the *Restatement (Second)* and *Restatement (Third)* treat this interpretation of the alternative liability rule as viable.[18]

In the next example, despite similarities to Example I.2, courts would clearly avoid any aggregation.

> *Example I.3. Injury in the Hospital: Two Events, Two Injuries*. Same facts as in Example I.2, except that it is alleged that each doctor negligently caused the plaintiff separate harm: the doctor in the emergency room allegedly caused him necrosis in his leg, and the doctor in the cardiology department allegedly injured his heart. The probability of each allegation is 40%.

---

16.   RESTATEMENT (SECOND) OF TORTS § 433B(3) (1965).

17.   RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM § 28(b) (2010).

18.   *See id.* § 28(b) cmt. k; *id.* § 28 reporters' note (citing several cases suggesting that the simultaneity of the defendants' acts is not a precondition for the applicability of the alternative liability rule); RESTATEMENT (SECOND) OF TORTS § 433B(3) cmt. h ("The cases thus far decided in which the rule stated in Subsection (3) has been applied all have been cases in which all of the actors involved have been joined as defendants. All of these cases have involved conduct simultaneous in time, or substantially so, and all of them have involved conduct of substantially the same character, creating substantially the same risk of harm, on the part of each actor. It is possible that cases may arise in which some modification of the rule stated may be necessary because of complications arising from the fact that one of the actors involved is not or cannot be joined as a defendant, or because of the effect of lapse of time, or because of substantial differences in the character of the conduct of the actors or the risks which they have created. Since such cases have not arisen, and the situations which might arise are difficult to forecast, no attempt is made to deal with such problems in this Section. The rule stated in Subsection (3) is not intended to preclude possible modification if such situations call for it."). Several courts, however, have rejected that application of the alternative liability rule. *See* Skipworth v. Lead Indus. Ass'n, 690 A.2d 169, 174 (Pa. 1997) (rejecting the argument that the alternative liability rule applies when defendants did not act simultaneously); Smith v. Cutter Biological, Inc., 823 P.2d 717, 725 (Haw. 1991) (same).

With no aggregation, the hospital—as well as the two doctors—will bear no liability, even though the probability that the plaintiff suffered harm caused by wrongdoing for which the hospital is responsible is 64%. Aggregation, instead, would lead to the imposition of liability on the hospital if the events are independent. Indeed, analogizing from the alternative liability rule of *Summers v. Tice*, one could argue that if the plaintiff could show that the two doctors were each negligent, each doctor should be held liable unless he can establish that his negligence did *not* cause the plaintiff's harm. The *Restatement (Third)* justifies the alternative liability rule in part by explaining that "between two culpable defendants and an innocent plaintiff, it is preferable to put the risk of error on the culpable defendants."[19] Although the rule does not apply, under current law, to two-event, two-injury cases, its underlying "risk-of-error" logic arguably applies to such cases, as well as to incidents of one event, one injury (as in *Summers v. Tice*) and two events, one injury (as in Example I.2).

If aggregation occurs in Example I.3, what amount of liability would be imposed on the hospital? At a minimum, the hospital would be liable for the less severe injury. Alternatively, the hospital could be liable for the average of the two injuries, for the more severe injury, for a probabilistic recovery, or for some other amount. Each of these options has both advantages and disadvantages that we discuss later.[20] At this stage, it suffices to say that nonaggregation in cases represented by both Examples I.2 and I.3 would allow defendants to escape liability even when the probability that they wrongfully harmed the plaintiff (or are vicariously liable for the harm) is greater than 50%, just because the plaintiff cannot identify with sufficient certainty the *exact* wrongful injurious behavior that caused his harm (Example I.2), or which of the two harms suffered by him is the result of an injurious behavior (Example I.3).[21]

So far, aggregating claims would lead to more rather than less liability. But this is not always so. In Example I.3, with different numbers, aggregation could lead to less rather than more liability. To see why, assume that the

---

19. RESTATEMENT (THIRD) OF TORTS § 28(b) cmt. g. The comment explains that "[i]n at least some cases, it appears that the defendants' better access to proof and doubts about the plaintiff's ability to extract that evidence from the defendants . . . have influenced the courts to employ burden shifting."

20. *See infra* Subsection V.C.3.

21. We might also imagine a one-event, two-injury (or two-harms-from-one-injury) case that would fall between Examples I.2 and I.3, in which a plaintiff establishes that he suffered two distinct injuries (or two separate harms from one injury) during a single event and attempts to hold one defendant liable. As in Examples I.2 and I.3, aggregating the probabilities that the defendant caused each injury (or each harm) could enhance liability in such a case.

probability of the claims with respect to each doctor is 60% instead of 40%. With no aggregation the hospital would be liable for both injuries; with aggregation, the hospital will be liable for one injury only. The probability that the claims against both doctors are correct is only 36%,[22] and 36% is not enough to establish liability. The probability that at least one of the claims is correct, however, is 84%,[23] more than enough to establish liability.

Note that less liability in this example is arguably consistent with tort law goals. Tort law seeks to impose liability if, and only if, it is more probable than not that the defendant wrongfully caused the plaintiff's injury. In the general case, it is probably assumed by the law that the point where the adverse effects of too much and too little liability (and compensation) equalize is where the probability of the plaintiff's allegations is 50%.[24] Under this assumption, if the probability that both injuries to the plaintiff were wrongfully caused by the two doctors is only 36%, liability for both harms should not be imposed.

The point that aggregating claims could lead to either more or less liability in cases involving two injuries brings us to an interesting conclusion. When an injurer cannot know in advance whether, in his case, aggregation would increase or reduce his liability, aggregation would not necessarily change the injurer's expected liability and so might not affect his behavior.

To illustrate, assume that in our example the hospital anticipates that there could be two injuries where in each case the harm would be 100, and the probability that a doctor's negligence caused each injury will be either 40% or 60% for both injuries (with equal probabilities). With *no* aggregation the hospital's expected liability if the two allegations are made is 100. There is a 50% likelihood that the probability of each allegation is 40% and the hospital pays zero, and a 50% likelihood that the probability of each allegation is 60% and the hospital pays 200. But the hospital's expected liability is also 100 *with* aggregation. If courts will aggregate, then there is still a 50% likelihood that the probability of each allegation is 40%, but now the hospital would pay 100, because there is a probability of 64% that a doctor was responsible for at least one of the two injuries. There is also a 50% likelihood that the probability of each allegation is 60%, but now the hospital would pay 100 here as well, because there is a probability of 84% that a doctor was responsible for at least

---

22. $0.6^2 = 0.36$.

23. $1 - (1 - 0.6)^2 = 0.84$.

24. For more on the logic of the preponderance-of-the-evidence rule and its role in allocating the risk of error, see ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 143-48 (2005). For a critical account of the argument that the preponderance-of-the-evidence rule maximizes social welfare, and for the argument that the efficient level of proof is contextual, see Louis Kaplow, *Burden of Proof*, 121 YALE L.J. 738 (2012).

one of the two injuries, but a probability of only 36% that a doctor was responsible for both injuries. Once expected liability with or without aggregation is the same, the injurer's incentives are the same as well.

Aggregation, however, would be of utmost importance for efficient incentives in two-injury cases if the injurer could know in advance that the probabilities of each of the claims against him would typically be lower than 50%, or would typically be higher than 50%. In the extreme case where the probabilities are always lower than 50%, with no aggregation the injurer never pays and is underdeterred, while with aggregation he pays sometimes and is better deterred.[25] Conversely, if the injurer could know in advance that the probabilities in his case would always be higher than 50%, a rule of no aggregation could result in overdeterrence,[26] because under the rule of no aggregation the injurer's expected liability would be higher than the expected harm of his behavior. Aggregation would reduce expected liability, making it closer to the expected harm and improving deterrence.

One type of situation where the injurer could know in advance whether the probabilities of liability would be lower or higher than 50% is a medical malpractice case, in which evidence regarding causation is likely to be statistical in nature. A doctor can often predict whether the probability of causation that a

---

25. Note, however, that in cases represented by Example I.3, under certain assumptions (which we believe very rarely hold), aggregation could have some overdeterrence effects that should be contrasted with the underdeterrence effects that nonaggregation yields. Thus, if the second doctor in this example is fully informed, in "real" time, about the first doctor's alleged negligent infliction of harm on the plaintiff, and also that the probability of this allegation is substantial but lower than 50%, he (or the hospital) would take excessive precautions so as not to bear liability if something goes wrong and the plaintiff suffers a second injury. An analogous argument could be made with respect to Example I.2 (Injury in the Hospital: Two Events, One Injury). In other circumstances, and under similar assumptions, aggregation may lead to underdeterrence. Imagine that in the prior example, the second doctor is informed in real time that the probability of the allegation against the first doctor is slightly *more* than 50%. The second doctor (or the hospital) might take deficient precautions to prevent a second injury for the same patient, knowing that only a very high-probability allegation with respect to such a second injury would allow him any additional recovery for it. (To illustrate, if the probability of the patient's allegation relating to the first injury is 51%, in order to recover for both the first and second injury, the patient should be able to prove his allegations with respect to the second injury with a probability close to 100%.)

26. Under a negligence rule, overdeterrence would arguably not result if the standard of care were set accurately, the injurer could observe it, and the courts could accurately enforce it, even if the injurer pays damages higher than the harm caused by his negligence. *See* ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 218-19 (6th ed. 2012) (arguing that a small change in the damages awarded to the victim will not cause the injurer's behavior to change).

plaintiff will be able to establish using statistical evidence is lower or higher than 50%.[27] In other types of cases the injurer may just know that it would be hard for the victim to collect evidence — statistical or case-specific — to prove his case, so he can predict that the probabilities of any allegations of injury would be low.

Lastly, there are cases where aggregation can lead to more liability, but never less. These are the cases where two (or more) alternative claims are made regarding a single injury, rather than two separate injuries. Examples I.1 (The Inherently Dangerous Vehicle) and I.2 (Injury in the Hospital: Two Events, One Injury) belong to this category. No aggregation in this context typically leads to underdeterrence, because in some cases injurers escape liability even if it is more probable than not that they wrongfully inflicted harm on the defendant. Aggregation can help to correct this underdeterrence.

In all of the examples discussed so far, aggregation of claims would not be necessary if courts allowed probabilistic recoveries. Under a probabilistic recovery rule (PRR), a defendant's liability is the amount of the harm done to the plaintiff multiplied by the probability that the harm was wrongfully caused by the defendant. Only some jurisdictions allow PRR, and even when they allow it, PRR applies in very limited contexts (mostly in cases of lost chances of recovery[28]) and only when causation — but not wrongfulness — is uncertain.[29]

It is beyond the scope of this Article to discuss the advantages and disadvantages of PRR. But some brief comments are in order. There are some clear limits to PRR compared to aggregation. First, when the remedy cannot be prorated, as with injunctions, PRR is inapplicable, while aggregation applies more broadly.[30] Second, to apply PRR, courts need accurate information about

---

27. For a discussion of medical malpractice cases where the typical probabilities of causation are low and underdeterrence may result, see Ariel Porat, *Misalignments in Tort Law*, 121 YALE L.J. 82, 108-14 (2011).

28. *See, e.g.*, Herskovits v. Grp. Health Coop. of Puget Sound, 664 P.2d 474, 476-77 (Wash. 1983) (holding that a fourteen percentage-point reduction, from 39% to 25%, in the decedent's chance for survival was sufficient evidence to allow the case to go to the jury); Joseph H. King, Jr., *Causation, Valuation, and Chance in Personal Injury Torts Involving Preexisting Conditions and Future Consequences*, 90 YALE L.J. 1353, 1394-95 (1981) (endorsing the application of a probabilistic rule to lost-chance-of-recovery cases); *see also* Doll v. Brown, 75 F.3d 1200, 1206-07 (7th Cir. 1996) (suggesting the use of a probabilistic rule to calculate damages for the lost chance to be promoted).

29. *See generally* ARIEL PORAT & ALEX STEIN, TORT LIABILITY UNDER UNCERTAINTY 57-83 (2001) (discussing the pros and cons of probabilistic recovery in different contexts, including cases where wrongfulness is uncertain).

30. To be fair, tailoring a remedy under aggregation could sometimes be complex. Thus, in Example I.3 (Injury in the Hospital: Two Events, Two Injuries), as we have explained, there

the probability that the defendant wronged the plaintiff. That could often make adjudication costly, which is probably one of the reasons why PRR is so rare in the law. This is not the case with aggregation. To aggregate claims, courts do not need to calculate exact probabilities; they only need to determine whether aggregation makes it more probable than not that the defendant wronged the plaintiff (or that a defense applies). Third, corrective justice theorists resist PRR because it can make the defendant compensate the plaintiff even though it is not more probable than not that he has wronged him. Other theorists not belonging to the corrective justice school could raise the related concern that the machinery of the law should not be put in motion against a defendant until it is more probable than not that the plaintiff is entitled to a remedy. In contrast, with aggregation, the defendant pays damages (or is subject to other remedies) only when it is more probable than not that he wronged the plaintiff.[31] Aggregation would therefore be easier to accept for many who resist PRR.[32]

At the same time, PRR holds several advantages over aggregation. In particular, PRR works more systematically and accurately than aggregation because it is applied to each claim separately and calibrates damages accurately. Also, possible strategic behavior by injurers trying to avoid future aggregation of low probability claims against them under an aggregation rule would not take place under PRR.

The last example in this Section, which was first analyzed by Saul Levmore,[33] addresses a case where factual aggregation occurs within a claim but across elements—which we call *cross-element factual aggregation*. Such aggregation typically leads to less rather than more liability.[34]

---

are several possible awards of damages under an aggregation rule, and it is often hard to know how to choose among them. For further discussion of the aggregation of disparate monetary awards, see *infra* Subsection V.C.3. Also, there could be situations where the plaintiff seeks to aggregate claims for two different types of relief, like damages and an injunction. Tailoring an aggregated remedy in such a case could be especially complex. *See infra* Subsection V.C.3.

31. The alternative liability rule is a narrow exception, as it permits liability where a defendant's wrongdoing, but not the causal link between that wrongdoing and the plaintiff's injury, has been established by the preponderance of the evidence. See *supra* notes 15-18 and accompanying text.

32. For more on the relationship between corrective justice theory and aggregation, see *infra* Subsection V.B.1.

33. Levmore, *supra* note 8, at 723, 725-28.

34. It could lead to more liability only if some elements were alternatives to one another. *Id.* at 726-29.

Example I.4. Several Elements of One Claim (A). In support of his negligence claim, the plaintiff argues that the defendant acted negligently and that the defendant's conduct caused the plaintiff's injury. The probability that the defendant acted negligently is 60%, and the probability that the defendant's conduct caused the plaintiff's harm is also 60%.

With no aggregation, the court would find the defendant liable, but with aggregation, liability would not be imposed, assuming the probabilities are independent. Specifically, aggregation would yield that the probability that the defendant negligently caused the litigated harm is only 36%, short of the preponderance-of-the-evidence threshold. Note that the aggregation problem becomes more acute as the number of elements composing the claim rises. Thus, if—in addition to the uncertainty with respect to negligence and causation—there is also uncertainty with respect to the plaintiff's harm, such that each of the three elements is proven with a probability of 60%, aggregation would yield a probability of 21.6% that the defendant negligently caused the litigated harm.[35] The law is not clear as to whether juries and judges should engage in cross-element aggregation: in several jurisdictions, jury instructions encourage such aggregation, while in other jurisdictions they discourage it.[36]

Throughout this Section, we have assumed that the probabilities of the two (or more) claims being valid are independent; that is, if one claim is valid, it does not affect the probability of the other claim being valid. However, this assumption does not cover all cases. Sometimes there is dependence between the probabilities, and then aggregation becomes more complex.[37] Thus, if the defendant allegedly engaged in two separate wrongful acts that caused two injuries (or injuries caused by two individuals for whom the defendant is vicariously liable, as in Example I.3 (Injury in the Hospital: Two Events, Two

---

35. $0.6^3 = 0.216$.

36. Levmore, *supra* note 8, at 752 n.58 (suggesting that no jurisdiction explicitly recognizes cross-element factual aggregation); *id.* at 725, 741 (arguing that jury instructions tend to be ambiguous in several states, implicitly allowing such aggregation); *see also* Ronald J. Allen & Sarah A. Jehl, *Burdens of Persuasion in Civil Cases: Algorithms v. Explanations*, 2003 MICH. ST. L. REV. 893, 897-902 (arguing that virtually no model jury instruction, state or federal, endorses cross-element factual aggregation); Alex Stein, *Of Two Wrongs That Make a Right: Two Paradoxes of the Evidence Law and Their Combined Economic Justification*, 79 TEX. L. REV. 1199, 1204 (2001) (providing examples of jury instructions in several jurisdictions that call for separate examination of the elements).

37. *Cf.* Levmore, *supra* note 8, at 726-28 (discussing the dependence problem mainly in cases similar to our Example I.4).

Injuries)), the invalidity of the claim that the defendant negligently caused the first injury could decrease the probability of the validity of the claim that the defendant negligently caused the second injury, and vice versa. This complication, however, does not necessarily preclude aggregation.

To illustrate, assume that claim *A*'s probability standing alone is 40%, and claim *B*'s probability standing alone is also 40%. Assume now that the probabilities of the two claims are somewhat dependent: if claim *A* is invalid, the probability of claim *B* being valid is only 20%. With no aggregation, both claims will be dismissed since a 40% probability (and certainly a 20% probability) is not enough to establish liability. With aggregation, however, some degree of liability would be imposed, since the probability that at least one claim holds is greater than 50%.[38] Because the probabilities of the two claims are dependent, the probability that at least one claim is valid is less than 64%, which would have been the result of aggregation if both claims (with 40% probability) had been independent. In the extreme case, the dependence between the probabilities of the two claims is full, which means that if claim *A* is invalid, claim *B* is also invalid and vice versa. With full dependence, aggregation becomes meaningless, since the probability that claim *A* (or claim *B*) holds is the same as the probability that at least one of those claims holds.[39]

B. Normative Aggregation and Mixed Aggregation

Consider the following example:

Example I.5. Insanity and Mitigation: Two "Almost Defenses." The defendant hit the plaintiff while driving his car at an unreasonable speed. The plaintiff was injured and later chose not to undergo an essential surgery that would have cured him completely. The plaintiff sues the defendant for negligently causing him the injury. The defendant raises two defenses: insanity on his part and failure to mitigate damages on the plaintiff's part. The court concludes that even though the defendant suffered from severe mental instability at the

---

38. The probability of claim *A* being invalid is 0.6, and the probability of claim *B* being invalid, conditional on the invalidity of claim *A*, is 0.8. Therefore, the probability of both claims being invalid is 0.6 * 0.8 = 0.48, and the probability of at least one claim being valid is 1 - 0.48 = 0.52.

39. Notice that updating the probabilities of claims *A* and *B* due to their dependence is not aggregation in our terms. With aggregation, the probabilities of the two claims are not changed; instead, aggregation derives the probability that at least one of the two claims is valid. For further elaboration on this distinction, see *infra* notes 101-102 and accompanying text.

time of the accident, his mental capacity had not been diminished to the point where the insanity defense applies.[40] The court also concludes that even though the plaintiff's failure to undergo the surgery was unreasonable for most people, the mitigation-of-damages defense does not apply, since tort law tolerates people's resistance to undergoing surgery.[41]

The court deciding the case would not aggregate the two defenses raised by the defendant, and would reject both of them.[42] We might criticize this stance by pointing out that a defendant with both "almost defenses" may seem less blameworthy than a defendant with only one. In a metaphoric way, we could say that a defendant whose justifications or excuses for wrongdoing reach a point of normative weight denoted as a should not face liability. That point can be reached with a single viable defense that provides the normative weight of a, but also by the accumulative normative weight of two "almost defenses" (assuming, for example, that the normative weight of each of the "almost defenses" is $\frac{1}{2}a$ or more).[43] Consider this argument from an economic perspective. We sometimes do not impose liability on mentally incompetent people because they are undeterrable, and we deny damages to plaintiffs who fail to mitigate in order to give them an incentive to mitigate. But we may also want to deny damages where the barely mentally competent person will be only barely responsive to them *and* the surgery-fearing victim will be

---

40. The insanity defense is quite limited under American tort law. *See* Dobbs, *supra* note 9, § 120 (stating that the general rule is that the mentally disabled are liable for negligence, subject to limited exceptions); *cf.* Breunig v. Am. Family Ins. Co., 173 N.W.2d 619 (Wis. 1970) (deciding that when a person commits an act as a result of a sudden onset of unforeseeable insanity comparable in its effect to certain physical impairments, liability will not be imposed).

41. *See* RICHARD A. EPSTEIN, TORTS 448 (1999) (noting that a plaintiff "should, and typically does, operate within a domain of reasonable choice that spares her from having to make unwanted life choices solely to minimize . . . [a defendant's] financial losses").

42. In *Boa v. San Francisco-Oakland Terminal Railways*, 187 P. 2 (Cal. 1920), the plaintiff sustained physical injury when she exited the defendant's streetcar. *Id.* at 3. The defendant argued both that the plaintiff was contributorily negligent and that she failed to mitigate damages by choosing an improper physician. *Id.* at 5. The trial court instructed the jury that the contributory negligence defense should be rejected unless it had been proved by the preponderance of the evidence. *Id.* at 6. The jury denied both defenses. Had the court instructed the jury to aggregate the claims, either factually or normatively, the jury might have reached a different decision.

43. Notice that this argument holds even if one of the defenses raised by the defendant does not exist in the legal system at all: thus, even if an insanity defense is absent from a certain legal system, an argument can be made that the defendant's insanity coupled with an additional "almost defense" could be sufficient to exempt the defendant from liability.

somewhat responsive to the absence of them, because their joint response may well be optimal if damages are not awarded. At the margin, the driver's incentives will be less affected if damages are awarded than the victim's incentives will be if damages are denied.

One possible solution is to aggregate the two "almost defenses" and release the defendant from liability for the harms that would have been avoided if the plaintiff had undergone the surgery. Like an aggregated damages award that grants a plaintiff the lesser of two remedies, this aggregated defense would afford the defendant the lesser of two reprieves from liability.[44] By doing so, the court would acknowledge that even if none of the defenses *standing alone* should apply, the cumulative weight of the two "almost defenses" is sufficient to justify a reduction of liability.

As we have said, courts would probably not allow aggregation in Example I.5. They might, however, be more receptive to aggregation arguments when facing two defense arguments based on similar normative grounds, such that the defenses interrelate with each other. Thus, if a defendant raises the defenses of self-defense and insanity in tandem, arguing that the somewhat excessive force she used to defend herself was caused by her deficient mental capacity at the time of the injury, the justification to aggregate the two "almost defenses" might make more sense to some courts.[45]

However, courts should be cautious with cross-claim normative aggregation, because the weight of an "almost defense," or an "almost claim," could be zero, and then there would be nothing to aggregate. For example,

---

44. A successful mitigation defense would have excused the defendant from liability for the harms that surgery would have prevented, while a successful insanity defense would have excused the defendant from all liability.

45. For an example from criminal law, consider *State v. Peterson*, 857 A.2d 1132 (Md. Ct. Spec. App. 2004), where the defendant was convicted of murder in the first degree of her husband, and subsequently brought a post-conviction claim of ineffective assistance of counsel. *Id.* at 1135. The defendant claimed that her trial counsel had erred by failing to present evidence of, or a defense based on, the defendant's alleged battered spouse syndrome. *Id.* The court agreed and ordered a new trial. A state appellate court affirmed that decision, concluding that the defendant's trial counsel should have introduced evidence of battered spouse syndrome to support a claim of imperfect self-defense, which could have mitigated the defendant's liability. *Id.* at 1154. The court's holding could be given two related interpretations: (1) it implicitly aggregated the two "almost" defenses of battered spouse syndrome and self-defense; or (2) it recognized in the imperfect self-defense doctrine an aggregation of the concerns underlying the defenses of self-defense and battered spouse syndrome. Under this second interpretation, the imperfect self-defense doctrine would itself embody a type of aggregation. We later argue that the unconscionability doctrine in contract law and several doctrines in constitutional law embody aggregation in a similar way. *See infra* notes 72-75, 129-137 and accompanying text.

suppose a driver hits a pedestrian and then subsequently crashes into the pedestrian's house. A court holds that each act was almost but not quite negligent—in both cases, the cost of precaution would have been (barely) more than the expected harm. When the claims are aggregated, it remains the case that the defendant should not be held negligent—because the *joint* cost of precaution would have been greater than the *joint* expected harm.

That conclusion might change if we adopted a different theory of negligence. If, for example, we believe that there is some moral blame in causing harm nonnegligently, but that the level of blame by itself is not enough to justify the law's intervention, then we might believe that intervention is warranted once there is more than one injury caused by the same defendant to the same plaintiff (or maybe even to different plaintiffs). From such a theory of torts, one could develop a possible justification for strict liability for ultrahazardous activities: even if an injurer's activity is efficient, the high intensity of creating risks to victims is a justification for imposing liability on him.[46]

Cross-claim normative aggregation can explain a puzzling legal rule: the common law duty to rescue someone whom you have (even nonnegligently) placed in danger. Thus, a person who nonnegligently shot her gun in the forest and hit the plaintiff, causing him to fall into a pool of water, must take reasonable measures to rescue him, even though other people do not have such a duty.[47] Tort theorists have struggled to justify this rule, given that, independently, nonnegligently causing harm and nonrescue do not give rise to tort liability. Isn't it true that nil plus nil is still nil?[48] A possible explanation is that tort law aggregates two claims, each of which has some normative weight but neither of which is alone sufficient to justify liability, so that once those

---

46. *Cf.* Gregory C. Keating, *Distributive and Corrective Justice in the Tort Law of Accidents*, 74 S. CAL. L. REV. 193, 203-19 (2000) (arguing that when an activity creates risks that are concentrated on a few individuals, imposing liability for the creation of those risks is often justified by fairness considerations even if the creation of those risks is reasonable).

47. *See* Maladona v. S. Pac. Transp. Co., 629 P.2d 1001, 1004 (Ariz. Ct. App. 1981) (deciding that when the defendant creates the danger, even if with no fault, he has a duty to rescue even if the plaintiff was contributorily negligent); DOBBS, *supra* note 9, § 316 (arguing that an exception to the no-duty-to-rescue principle applies when the defendant who failed to rescue caused the harm or created a risk to the plaintiff, even if innocently and without fault); *see also* EPSTEIN, *supra* note 41, at 291 (giving an example of a car that blocks the highway without fault and explaining that the driver still has the responsibility to warn other drivers of the danger).

48. *See* Richard A. Epstein, *A Theory of Strict Liability*, 2 J. LEGAL STUD. 151, 193 (1973) (noting this problem and arguing that judges dislike the outcome of no liability in such cases, which explains this otherwise-unexplained exception to the no-duty-to-rescue rule). Following Epstein's logic, one could make the argument that the creation of the exception is the result of implicit normative aggregation.

two claims are aggregated, liability is justified. In particular, nonnegligently causing harm is not sufficient to justify liability, and failing to rescue is not sufficient to justify liability, but nonnegligently causing harm followed by a failure to rescue may nevertheless justify liability.[49]

If cross-claim factual aggregations and cross-claim normative aggregations were recognized, the door would be open for mixed aggregations. The next example, which is a variation of Example I.5, illustrates the potential for mixed aggregations.

> *Example I.6. Insanity and Mitigation: An Uncertain Defense with a Certain "Almost Defense."* Same facts as in Example I.5, except that there is factual uncertainty rather than normative insufficiency as to the application of the insanity defense. The probability that the defense applies is 40%.

With no aggregation, the court would reject the insanity defense in Example I.6, since the defendant failed to establish that defense by the preponderance of the evidence. The court would also reject the failure-to-mitigate defense, since the failure of the plaintiff to undergo surgery, even if considered unreasonable by most people, does not trigger the application of the defense.[50] Conversely, by aggregating the defenses, the court would conclude that some reduction in liability is appropriate. The court might then accept the

---

49. This is not the only possible explanation, of course. *See* EPSTEIN, *supra* note 41, at 291 (arguing that in cases where the defendant caused the injury with no fault, it becomes easier to identify the one who could have rescued the plaintiff); William M. Landes & Richard A. Posner, *Salvors, Finders, Good Samaritans, and Other Rescuers: An Economic Study of Law and Altruism*, 7 J. LEGAL STUD. 83, 125-26 (1978) (arguing that when the risk is increased by the actor, the risk of error in establishing causation between the omission and the injury is reduced); Ernest J. Weinrib, *The Case for a Duty To Rescue*, 90 YALE L.J. 247, 257-58 (1980) (asserting that the increase in the probability of an accident diminishes the ability of the victim and others to abate it, and therefore the defendant "must account for the increased risk").

50. In *Davenport v. F.B. Dubach Lumber Co.*, 36 So. 812 (La. 1904), the plaintiff, the defendant's employee, was run over by a locomotive and sustained severe injuries. *Id.* at 812. The defendant argued in its defense that the plaintiff was contributorily negligent, and that while being hospitalized he signed a compromise agreement with the defendant, releasing it from liability. *Id.* at 813-14. The plaintiff argued that he was not contributorily negligent and that when he signed the compromise agreement he was under the influence of drugs. *Id.* The court denied the defenses and decided for the plaintiff. From the facts in the court decision, it seems that there was some factual uncertainty as to the applicability of the contributory negligence defense. *Id.* If the defense relating to the release agreement was "almost" applicable, a mixed aggregation could have brought the court to a different decision.

lesser of the two independent defenses, exonerating the defendant from liability for the harm the plaintiff could have mitigated if she had undergone the surgery.[51]

Finally, a *cross-element normative aggregation* could be an interesting option to consider. Under this type of aggregation, when one element of a cause of action is more than satisfied, it could *sometimes* compensate for another element of the same cause of action that is not quite satisfied. Consider a variation of Example I.4:

> *Example I.7. Several Elements of One Claim (B)*. The plaintiff argues that the defendant was negligent and that the defendant's negligence is the cause of the plaintiff's injury. The defendant admits his negligence, but argues that the plaintiff suffers only stand-alone emotional harm, and therefore liability should not be imposed.

If the harm is indeed stand-alone emotional harm, then under current law, liability would most likely not be imposed: without an accompanying physical injury, courts usually consider the element of "harm" lacking. But cross-element normative aggregation could change the outcome. Assume that the defendant was not "just" negligent but grossly negligent, so that the element of "negligence" is more than satisfied. The defendant's gross negligence could compensate for the fact that the plaintiff suffers only stand-alone emotional harm, and could justify the imposition of liability. This type of aggregation would arguably comport with the goals of tort law. Courts may hesitate to impose liability for negligent infliction of stand-alone emotional harm because, in cases of ordinary negligence, such liability would lead to overdeterrence of injurers and unjustified litigation costs. Both concerns seem less compelling when the injurer's degree of fault is especially high. This way of reasoning

---

51. As with factual aggregation, the probabilistic recovery rule, or a version of it, could be applied in normative aggregation cases, making aggregation unnecessary. But as we have explained, there are several clear advantages to aggregation over the probabilistic recovery rule. *See supra* text accompanying notes 30-32. Furthermore, quantifying what percentage of a normative threshold each "almost claim" reaches — which would be necessary for applying a probabilistic recovery rule to normative aggregation cases — would be a different, and more foreign, task for courts than calculating probabilities of particular factual allegations. For a recent discussion of the related topic of the either-or character of the law, as opposed to probabilistic or other intermediate approaches, see LEO KATZ, WHY THE LAW IS SO PERVERSE 139-81 (2011), which also cites the main sources in the literature.

could explain why liability for stand-alone emotional harm is imposed under current law if the behavior of the defendant was outrageous.[52]

C. Cross-Person Aggregation

Should tort law allow cross-person aggregation? Take the following example:

> *Example I.8. Mass Torts: Indeterminate Plaintiffs.* The defendant's factory wrongfully emits radiation which causes an increase in the frequency of a fatal cancer in the population; instead of 100 people contracting the disease every year, now 125 people contract it every year. Because of a lack of scientific knowledge, it is impossible to identify those 25 victims whose disease was caused by the radiation. All 125 people bring suits against the defendant.[53]

Under traditional causation principles, all suits would be dismissed because none of the plaintiffs can establish by the preponderance of the evidence that her disease was caused by the defendant's wrongdoing. The plaintiffs can establish, however, that the wrongdoing caused harm to 25 out of the 125 plaintiffs. By aggregating all claims, the court could impose liability on the defendant for 20% of the total harm suffered by all plaintiffs, and could then distribute the damages among them.

Notice that the aggregation of all claims in Example I.8 mimics the defendant's liability if there were no uncertainty about which plaintiffs had been harmed by the radiation, but does not mimic the plaintiffs' entitlements if there were no uncertainty. Without uncertainty, the defendant would also have paid about 20% of the total harm, but those damages would have gone to 25 rather than 125 plaintiffs.

Market share liability, in contrast, mimics both defendants' liability and plaintiffs' entitlements as if there were no factual uncertainty. In the notorious diethylstilbestrol (DES) cases, numerous manufacturers produced the same

---

52. *See* DOBBS, *supra* note 9, §§ 302-04 (discussing the rule against liability for stand-alone emotional harm and its exceptions, principally where a defendant intentionally inflicted emotional distress).

53. For a well-known case presenting a similar problem, and in which a settlement was reached, see *In re "Agent Orange" Product Liability Litigation*, 597 F. Supp. 740 (E.D.N.Y. 1984), *aff'd*, 818 F.2d 145 (2d Cir. 1987). Veterans brought suits against manufacturers of "Agent Orange" for injuries allegedly caused by their exposure to that chemical, which was used by U.S. military forces in the Vietnam War. *Id.* at 746-47. Many of the plaintiffs' injuries could have been attributed either to their exposure to Agent Orange or to other causes. *Id.* at 777.

generic drug for preventing miscarriage. Many years later, the drug was found to be defective and harmful to the daughters of the women who had taken the drug.[54] Plaintiffs, however, found it impossible to prove the identity of the specific manufacturer that had produced the specific drug taken by their mothers many years earlier.[55] For some time, courts had refused to impose liability on manufacturers, since the probability that a specific manufacturer had actually caused the litigated harm in any given case was much lower than 50%.[56] In 1980, however, the California Supreme Court established the market share liability doctrine, whereby all manufacturers were held liable toward plaintiffs in accordance with their market share at the time a plaintiff's mother took the drug.[57] Under market share liability, when all suits are completed, manufacturers should bear liability in the amount of the actual harm they wrongfully caused and plaintiffs should receive damages in the amount of the harms they suffered from wrongdoing. Thus, market share liability aims at mimicking the world without uncertainty.

So far, we have illustrated cross-person aggregation under factual uncertainty. Thus, the cross-person aggregation discussed so far is an extension of the cross-claim factual aggregation discussed above in Section I.A. But cross-person aggregation can also extend to the cross-claim normative aggregation and cross-claim mixed aggregation discussed in Section I.B.

In *United States v. Hatahley*,[58] Native Americans brought suit against the U.S. government for trespass, arguing that their horses and burros were unlawfully rounded up by the government's agents and later sold to a horse-meat plant and a glue factory. Among other things, they sued for mental pain and suffering. The district court awarded them damages for mental pain and suffering under a theory that the emotional harm they suffered was "a community loss and a community sorrow shared by all."[59] That theory allowed

---

54. *See* Sindell v. Abbott Labs., 607 P.2d 924 (Cal. 1980).

55. *See id.* at 928.

56. *See, e.g.*, McCreery v. Eli Lilly & Co., 150 Cal. Rptr. 730, 733-35 (Ct. App. 1978) (denying liability of diethylstilbestrol (DES) producers and ruling that recovery for injuries resulting from a defective product requires that the plaintiff identify the manufacturer and establish the causal relation between the injury and the product).

57. *Sindell*, 607 P.2d at 937-38. The court defined the relevant "market" as sales of DES for the prevention of miscarriage. *Id.*

58. 257 F.2d 920 (10th Cir. 1958).

59. *Id.* at 925 n.5 (quoting the district court opinion). The district court further explained, as quoted by the appellate court:

> It is not possible for the extent of the mental pain and suffering to be separately evaluated as to each individual plaintiff. It is evident that each and all of the

the court to be generous to the plaintiffs and award them relatively high damages. The appellate court rejected the district court's theory, maintaining that "pain and suffering is a personal and individual matter, not a common injury, and must so be treated."[60] Using our terminology, the district court arguably allowed a kind of cross-person normative aggregation, perhaps under the assumption that the aggregate harm across persons exceeded a normative threshold even if the harm caused to any particular person did not.[61]

II. CONTRACT LAW

A. Factual Aggregation

Contract law, like tort law, appears to permit factual aggregation within elements,[62] but it does not generally permit cross-claim factual aggregation.

> *Example II.1. Either Material Breach or Fraudulent Misrepresentation.* The plaintiff can prove with 40% probability that the defendant engaged in a material breach of a contract, and also can prove with 40% probability that the defendant engaged in fraudulent misrepresentation. Under either theory, the plaintiff would be entitled to rescission of the contract.

Although the probability that at least one claim is valid is 64%, the plaintiff would lose because courts do not permit cross-claim factual aggregation.[63]

> plaintiffs sustained mental pain and suffering. Nor is it possible to say that the plaintiff who lost one or two horses sustained less mental pain and suffering than plaintiffs who lost a dozen horses.

Id.

60. *Id.* at 925. The court reversed and remanded the case for a new trial as to damages. *Id.* at 926.

61. Alternatively, the district court's claim that "[i]t is not possible for the extent of the mental pain and suffering to be separately evaluated as to each individual plaintiff," *id.* at 925 n.5 (quoting district court opinion), might suggest that the district court engaged in cross-person *factual* aggregation. Perhaps separate evaluations were not possible because mental pain and suffering could not be demonstrated with sufficient certainty on an individual basis. It is also possible that both cross-person normative aggregation and cross-person factual aggregation played a role at the same time, which may imply that the district court decision exhibited cross-person mixed aggregation.

62. We have not found any cases that clearly recognize (or reject) within-element factual aggregation in contract law, but we would be surprised if there were a major difference between the way contract law and tort law address such aggregation.

63. Or so we believe. We have not found any cases suggesting otherwise.

Long-term contracts or business relationships involving multiple contracts can raise issues of cross-claim factual aggregation. In such cases, a buyer might bring a suit against a seller arguing that the seller breached the same contract several times in the past, or breached several contracts in the past, and therefore owes the buyer compensation. The buyer might fail, however, to prove any specific breach by the preponderance of the evidence, so the court, without aggregation, would dismiss the suit. By contrast, aggregation would lead the court to award damages for one or several of the alleged breaches; or, in the appropriate case, to endorse the buyer's refusal to offer payments for performance because of one, or more, of the alleged breaches. In awarding relief for such aggregated wrongdoing, the court would not be able to point out the exact breach that took place. Accordingly, like a court that aggregates tort claims with different damages awards,[64] the court here would have to choose an alternative means of calculating liability. The court could, for example, craft a remedy that averaged the alleged breaches. If the alleged breaches are similar (say, five deliveries of the same number of widgets, with similar allegations of breach), arithmetic averaging would be relatively easy. Factual discrepancies between the alleged breaches, however, would render aggregation more complex, and one could argue that aggregation should not be made by averaging where the breaches differ substantially. In such cases, the court could still aggregate by acknowledging, at a minimum, that the least severe breach took place and allowing a remedy for it.

Parties to contracts, unlike tort victims and wrongdoers, can address aggregation directly by providing in their contracts that the court should aggregate facts. As far as we are aware, parties do not do so.[65] That raises the question whether cross-claim factual aggregation is actually a desirable approach. It may be that parties do not provide for cross-claim factual aggregation because it would not improve incentives. Although cross-claim factual aggregation leads to more accurate decisions ex post, it would not improve incentives if the too-high and too-low liability outcomes, which aggregation would correct, cancel out ex ante.[66]

---

64. *See supra* notes 20-21 and accompanying text (discussing analogous remedial challenges with respect to Example I.3 (Injury in the Hospital: Two Events, Two Injuries)).

65. At least not explicitly. It is possible, however, that parties do open the door for aggregation in more subtle ways. For instance, contracts often call for cooperation, such as best efforts or good faith — relatively amorphous standards to which several distinct acts might be relevant. Contracts also often create mechanisms for the resolution of disagreements by nonlawyer arbiters who need not provide rigorous and formal reasoning for their decisions. Those standards and mechanisms could be used for implicit aggregations.

66. *See supra* p. 15 (discussing cases in which aggregation in tort would not change a potential injurer's incentives).

To illustrate, assume that the parties anticipate, when making their contract, that there could be two allegations of two separate breaches by the promisee, where in each case the harm would be 100 and the probability of each breach would be either 40% or 60% for both breaches (with equal probabilities). With *no* aggregation the promisor's expected liability if the two allegations are made is 100: there is a 50% likelihood that the probability is 40% and he pays zero, and a 50% likelihood that the probability is 60% and he pays 200. But the promisor's expected liability is also 100 *with* aggregation. If courts aggregate, there is still a 50% likelihood that the probability is 40%, but now the promisor would pay 100, because there is a probability of 64% that he committed at least one of the two breaches. There is also a 50% likelihood that the probability is 60%, but now the promisor would pay 100 here as well, because there is a probability of 84% that he committed at least one of the two breaches, but a probability of only 36% that he committed both. Once expected liability with or without aggregation is the same, the parties' incentives are the same as well.[67]

If, however, the parties to the contract anticipate rescission as the remedy, and the two alleged breaches are material (or there is an alleged material breach and an alleged fraudulent misrepresentation, as in Example II.1), aggregation could only increase the promisor's ability to rescind the contract, and never decrease it.[68] That would generally improve the parties' incentives by correcting underdeterrence, because the promisee would be able to rescind the contract whenever it is more probable than not that the promisor materially breached it (or, in Example II.1, either materially breached the contract or engaged in fraudulent misrepresentation, or did both).

B. Normative Aggregation

Contract law does not directly permit cross-claim normative aggregation of the following type:

Example II.2. Non-Material Breach and Misleading Conduct. The defendant engages in misleading conduct that falls short of fraudulent

---

67. Under certain conditions, aggregation could lead to less liability than no aggregation. That does not change, in substance, the arguments favoring aggregation. *See supra* notes 22-24 and accompanying text (addressing this issue in the tort context). Also, aggregation could sometimes, under certain assumptions, lead to either overinvestment or underinvestment in precautions in preventing the second breach. *See supra* note 25 and accompanying text (addressing analogous tort issue).

68. *See supra* p. 17 (making an analogous point regarding tort cases involving two alternative claims stemming from a single injury).

misrepresentation in order to secure the plaintiff's consent to a contract. Subsequently, the defendant engages in a breach that falls just short of material. The plaintiff seeks to rescind the contract based on claims of fraudulent misrepresentation and material breach.

A court would typically not approve the rescission in Example II.2.[69] The plaintiff would lose on the first claim because the defendant's misleading conduct at the time of contract formation would fall short of fraudulent misrepresentation, and on the second claim because the breach is not material. Therefore, the court would decide that the plaintiff, by unlawfully rescinding the contract, breached the contract himself. By contrast, cross-claim normative aggregation would permit the plaintiff to claim that he was entitled to rescind the contract on the basis of both the fraudulent misrepresentation and the breach, even though neither of them standing alone was sufficient for rescission.

Yet contract law often seems to permit cross-element normative aggregation, a notable distinction from how tort law approaches aggregation. Consider the following example:

Example II.3. Two Minor Breaches. The defendant promises to build a house for the plaintiff. When the time for the first progress payment comes around, the defendant is a little behind in schedule and has made some minor mistakes in construction. The plaintiff claims that the defendant has breached the contract.

Even though each of the two breaches might not be regarded as substantial individually, a court could find them collectively substantial, justifying rescission on the part of the plaintiff. Courts often perform this type of normative aggregation to find contractual breach.[70] Essentially, each individual

---

69. At least not expressly. Courts sometimes permit considerations from one claim (or defense) to bleed over to another. For example, in *Lincoln Benefit Life Co. v. Edwards*, 45 F. Supp. 2d 722, 746-51 (D. Neb. 1999), a court found that the plaintiff had entered into a contract under duress in part because the defendant had also committed fraud by lying about the contents of the contract.

70. *See* SunTrust Mortg., Inc. v. United Guar. Residential Ins. Co., 806 F. Supp. 2d 872, 902 n.64 (E.D. Va. 2011) ("Precedent shows that, in assessing the materiality of multiple breaches, as the Court must do here, it is appropriate to consider the combined – or 'cumulative' – effect of the breaches."); *see also* Seven-Up Bottling Co. (Bangkok) v. PepsiCo Inc., 686 F. Supp. 1015, 1024 (S.D.N.Y. 1988) (concluding that the plaintiff's multiple breaches "with respect to minimum sales and distribution requirements" collectively constituted "a breach of its further obligation . . . to adequately promote and develop the market for [the defendant's] products").

breach functions as an element of the plaintiff's overarching claim that the defendant has violated the contract to the point of justifying rescission.[71] Consider another example, which shows the evolution of contract law to endorse a similar type of normative aggregation:

> *Example II.4. Unconscionability*. A store sells a TV set on credit to a poor customer. The customer is not well educated and does not read the contract, which provides that the store may repossess all of the goods that the customer bought previously from the store on credit if they are not yet fully paid for and the customer defaults on payments for the TV set.

Under older doctrine, if the customer defaulted and lost all goods previously purchased, she would not have a remedy. If she sued under the mistake doctrine, she would lose because she did not read the contract. If she sued on the grounds that she was uneducated, she would lose because although courts recognize incompetence or undue influence as grounds for rescission, they generally treat lack of education as falling short of incompetence or undue influence.[72] But over the last half century, the doctrine of unconscionability evolved.[73] Under this doctrine, a plaintiff can invoke considerations jointly — like failure to read the contract and lack of education — that can be considered only individually under other doctrines. In this way, the unconscionability

71. Courts have also arguably treated individual breaches as elements of the claim that a defendant violated the covenant of good faith and fair dealing, and have aggregated those breaches to impose liability. *See, e.g.*, Photovest Corp. v. Fotomat Corp., 606 F.2d 704, 727-30 (7th Cir. 1979) (holding that multiple breaches by a franchisor amounted to a violation of the covenant of good faith and fair dealing, and even a possible tort claim justifying punitive damages under Indiana law).

72. *See, e.g.*, Mason v. Acceptance Loan Co., 850 So. 2d 289, 296, 302 n.9 (Ala. 2002) ("The plaintiffs also argue that their poor reading skills and lack of education, and the defendants' alleged knowledge of those limitations, render the arbitration agreement unconscionable. However . . . we have repeatedly held to the contrary.").

73. *See* Williams v. Walker-Thomas Furniture Co., 350 F.2d 445 (D.C. Cir. 1965) (recognizing unconscionability to include an absence of meaningful choice caused by the inequality of bargaining power when one of the parties is uneducated and signed the contract without full knowledge of its terms); *see also* U.C.C. § 2-302(1) (2010) ("If the court as a matter of law finds the contract or any clause of the contract to have been unconscionable at the time it was made the court may refuse to enforce the contract, or it may enforce the remainder of the contract without the unconscionable clause, or it may so limit the application of any unconscionable clause as to avoid any unconscionable result."). *See generally* E. ALLAN FARNSWORTH, CONTRACTS 307-16 (3d ed. 1999) (discussing the unconscionability doctrine); JAMES J. WHITE & ROBERT S. SUMMERS, UNIFORM COMMERCIAL CODE 151-77 (2000) (same).

doctrine can be understood as permitting a kind of cross-element normative aggregation. Unconscionability is the claim, and issues like the plaintiff's failure to read and lack of education become elements of that claim. But it is important to see that the cross-claim normative aggregation is implicit in the design of the unconscionability cause of action: courts do not say that plaintiffs can prevail by presenting colorable claims under two doctrines, but the unconscionability doctrine has the same effect.[74]

As this example shows, courts may address concerns about aggregation through doctrinal evolution. Broader standards subsume narrower rules as courts realize that cases can fall between the rules even though the cases raise the concerns that justify those rules. But as the doctrine becomes broader and permits greater aggregation, critics complain that the law becomes too vague and can no longer guide behavior.[75] We will return to this problem in Part V.

C. Cross-Person Aggregation

As with torts, there are cases in contracts where plaintiffs cannot prevail against a specific defendant because of inherent difficulties of proof. These cases raise questions about whether cross-person aggregation should be allowed. The next example illustrates such a case.

> *Example II.5. Many Unproven Breaches with Customers.* The defendant ships goods by sea, and the plaintiffs are the defendant's customers whose goods were damaged. In most cases there is evidence indicating that the damage could be the result of a breach of contract by the defendant, but that evidence is not strong enough to establish liability in any one plaintiff's suit.

---

74. A similar point can be made about the doctrine of undue influence as applied in *Odorizzi v. Bloomfield School District*, 54 Cal. Rptr. 533 (Dist. Ct. App. 1966). In that case, the plaintiff sought rescission of his resignation from the defendant employer under several theories, including "duress, menace, fraud," "mistake," and "undue influence." *Id.* at 538. The district court sustained the defendant's demurrer regarding all of the plaintiff's claims. *Id.* While the appellate court also rejected his independent duress, menace, fraud, and mistake claims, it concluded that the plaintiff *had* stated a claim for undue influence. *Id.* at 539-43. In reaching this conclusion, the court arguably took into account and aggregated factors relevant to the other claims it had denied. *Id.*

75. For this criticism of the unconscionability doctrine, see, for example, Evelyn L. Brown, *The Uncertainty of U.C.C. Section 2-302: Why Unconscionability Has Become a Relic*, 105 COM. L.J. 287, 288 (2000), which argues that the unconscionability doctrine allows courts such "wide latitude" that they often "manipulate the unconscionability principle in order to reach the equitable results they desire."

The court in Example II.5 would reject all suits brought independently by individual plaintiffs because those plaintiffs would not be able to establish liability. Prevailing law would also not allow recovery even if the plaintiffs brought a class action, since any named plaintiff would still need to establish his or her individual injury.[76]

That result would change if courts were willing to aggregate all claims and allow full recovery in some of the cases, or partial (probabilistic) recovery in all of the cases. We speculate that the argument for aggregation would seem more compelling for courts if a breach were already established by the preponderance of the evidence, and only causation could not be proven without aggregation, than if the breach itself could not be proven without aggregation. As the development of the alternative liability rule demonstrates, courts applying tort law have been more generous to plaintiffs on the issue of causation when those plaintiffs have persuasively established a defendant's wrongdoing.[77] Under the alternative liability rule, once a plaintiff proves that a defendant acted wrongfully, courts are willing to place the risk of error regarding causation on the defendant.[78] We suspect that courts applying contract law would have a similar impulse, and so would more readily allow aggregation to establish causation when a plaintiff had already proven the defendant's breach without aggregation.

III. CRIMINAL LAW

A. Factual Aggregation

Aggregation in criminal law resembles aggregation in tort law but raises special concerns because of sensitivities about the rights of the accused.[79] Consider the following example:

---

76. *See* Simon v. E. Ky. Welfare Rights Org., 426 U.S. 26, 40 n.20 (1976) ("[E]ven named plaintiffs who represent a class 'must allege and show that they personally have been injured, not that injury has been suffered by other, unidentified members of the class to which they belong and which they purport to represent.'" (quoting Warth v. Seldin, 422 U.S. 490, 502 (1975))).

77. *See supra* notes 15-18 and accompanying text (discussing the alternative liability rule). In cases where courts allowed cross-person aggregation in tort law, defendants' wrongdoing had also already been established. *See supra* Section I.C.

78. *See supra* note 19 and accompanying text (explaining the "risk-of-error" logic behind the alternative liability rule).

79. *See* Harel & Porat, *supra* note 8 (discussing factual aggregation in criminal law).

Example III.1. Two Unproven Charges. The defendant is charged with pickpocketing and rape, two unrelated offenses allegedly committed by him at different times and places. The evidence suggests that the probability that he committed each one of these offenses is 90%. Assume that the required probability necessary to satisfy the beyond-a-reasonable-doubt standard is 95%.[80]

Under prevailing law, the defendant would be acquitted of both offenses. Yet there is a 99% probability[81] that he committed at least one offense, which is higher than the 95% probability necessary for conviction in a criminal trial. If instead the court engaged in cross-claim factual aggregation, it could convict the defendant of one unspecified offense and impose on him, at a minimum, the sanction of the less severe of the two offenses, pickpocketing.[82] Example III.1 raises a straightforward dilemma: individuals are routinely convicted for committing a single offense on the basis of evidence that establishes guilt with a lower probability (95% under our assumption) than the probability that the defendant in Example III.1 committed at least one offense (99%). Arguably, it is not just that the Example III.1 defendant is acquitted while, at the same time, a defendant charged with a single offense that can be proven at a lower probability (95% under our initial assumption) is convicted.

Example III.1 illustrates how cross-claim factual aggregation could result in more convictions than with no aggregation. But aggregation could also result in fewer convictions, as is illustrated in Example III.2.

Example III.2. Two Proven Charges. The defendant is charged with pickpocketing and rape, two unrelated offenses, allegedly committed by him in different times and places. The evidence suggests that the probability that he committed each one of these offenses is 95%. Assume that the required probability necessary to satisfy the beyond-a-reasonable-doubt standard is 95%.

Under prevailing law, the defendant would be convicted of both charges because the probability that he committed each of the offenses (95%) is

---

80. This example is borrowed from Harel and Porat, *id.* at 262.

81. The probability that the defendant committed at least one of the offenses is $1 - 0.1^2 = 0.99$.

82. Cross-claim factual aggregation could also apply to affirmative defense claims in criminal actions. For example, in *Ralston v. State*, 927 N.E.2d 430 (Ind. Ct. App. 2010) (unpublished memorandum decision), the defendant raised two defenses: that he did not cause the victim's injuries and that he acted in self-defense when he repeatedly punched the victim. Each claim was considered separately by the jury and denied. Aggregating the two claims might have brought a different result.

sufficient for conviction. Yet, the probability that the defendant committed *both* offenses is only 90%, which is lower than 95%.[83] Therefore, with cross-claim factual aggregation, the court would convict the defendant of only one offense: while the probability that he committed at least one offense is greater than 95%, which is sufficient for conviction, the probability that he committed two offenses is insufficient. The court would then need to decide which offense to convict the defendant of. In our view, the correct decision would be to convict the defendant only of the more severe offense, since the probability that he committed that offense is 95%.[84] However, it could well be appropriate for the court to nudge the sentence for the more severe offense up a bit to reflect the fact that there is a 90% probability that the defendant committed both offenses.[85]

Cross-claim factual aggregation in cases illustrated by Example III.1 would arguably improve deterrence (assuming, as we must, that the beyond-a-reasonable-doubt standard should be taken as fixed). Under current law, when defendants are charged with several offenses, and the probability that they committed at least one of those offenses is very high, they still often escape conviction just because no *specified* offense can be attributed to them with sufficient certainty. Those defendants could be considered underdeterred under current law, and would then be better deterred with cross-claim factual aggregation. In contrast, cross-claim factual aggregation in cases illustrated by Example III.2 would reduce false convictions (that is, convictions of the innocent). Under current law, defendants who are charged with several offenses are often convicted of all of those offenses, even if the aggregated probability that they committed all of them is too low to meet the beyond-a-reasonable-doubt threshold. With cross-claim factual aggregation, those defendants will be convicted of fewer offenses, and many false convictions will be avoided.[86]

---

83. $0.95^2 \approx 0.90$.

84.   This is a critical distinction from Example III.1, where we recommended imposing the *less* severe penalty as an aggregated remedy because the probability that the defendant deserved the sentence for the more severe offense was lower than 95% (it was 90%).

85.   *Cf.* Talia Fisher, *Conviction Without Conviction*, 96 MINN. L. REV. 833 (2012) (calling for calibration of the sanctions imposed on defendants to the level of the uncertainty of their guilt, and arguing for the relaxation of the beyond-a-reasonable-doubt standard in appropriate cases).

86.   Under certain conditions, reducing the number of false convictions by increasing the burden of proof—the result of aggregation in cases illustrated by Example III.2—increases deterrence since it increases the difference between the expected sanctions of the guilty and innocent. *Cf.* Chris William Sanchirico, *Character Evidence and the Object of Trial*, 101 COLUM. L. REV. 1227, 1263-64, 1277-80 (2001) (suggesting that the admission of certain

Cross-claim factual aggregation in criminal law might be regarded as more objectionable than in torts and contracts, because of the concern that it would curtail the accused's rights (although, as illustrated by Example III.2, aggregation could also favor the accused). Allowing aggregation would require changes in procedure that many would consider undesirable. In particular, aggregation would require that the prosecution be allowed to bring several charges of different natures against the accused at the same trial, since it is hard to imagine that aggregation could take place if each charge were brought before a different jury or judge.[87] Aggregation could increase the burden on the defense, since defending against several charges, even if each has a low probability, could be harder and more costly than defending against one high-probability charge. Aggregation could also encourage abuse and strategic behavior by the prosecution because it is typically easier—maybe too easy—to bring many low-probability charges against the defendant than to bring one high-probability charge against him. (From a different perspective, however, bringing several low-probability charges together could reduce the expense of criminal litigation by economizing on enforcement costs.)

Furthermore, prosecutors might strategically avoid bringing two high-probability charges together (as in Example III.2) if they believe that they have enough evidence to establish guilt for each of the charges, out of fear that the court would aggregate the probabilities and convict the defendant of only one charge. This concern could be mitigated, however, if the accused were allowed to force the prosecution to bring the two charges in a single trial.[88]

Finally, a more substantive objection to aggregation could be that it would dilute the expressive function of criminal law. Thus, in Example III.1, with aggregation, the accused would be convicted of being either a rapist or a pickpocket, and his criminal record might literally list his offense as "rape or larceny." Some commentators may consider that outcome intolerable: punishing a person for an offense that the person may or may not have committed rather than for the offense that the person actually committed dilutes the expressive, educational, and communicative messages of

---

types of propensity evidence might *decrease* deterrence by *decreasing* the difference between the expected sanctions of the guilty and innocent).

87. Rule 8(a) of the Federal Rules of Criminal Procedure provides that two offenses may be joined in the same indictment if they "are of the same or similar character, or are based on the same act or transaction, or are connected with or constitute parts of a common scheme or plan." FED. R. CRIM. P. 8(a). Under Rule 14(a) of the Federal Rules of Criminal Procedure, courts may order separate trials if the joinder of offenses appears to prejudice a defendant or the government. *Id.* R. 14(a).

88. *See infra* Subsection V.C.1.

punishment. But, on the other hand, it is hard to see why acquittal of both offenses would be preferable, even from an expressivist point of view.[89]

Cross-element aggregation is also an issue in criminal law.[90] If several elements of the same offense must be proven to establish the defendant's guilt, then cross-element aggregation could generate a different outcome than if each element were considered separately. For instance, if convicting a person for burglary requires both trespass and intent to commit a crime, it is possible that even if each element of the offense (trespass and intent) can be proven beyond a reasonable doubt, reasonable doubt could still exist with respect to the cumulative presence of the two elements. Will the court convict the defendant under such circumstances? The answer is unclear.[91] As in tort and contract law, however, within-element factual aggregation is permissible and appears routine.[92]

Cross-claim factual aggregation should be distinguished from two existing doctrines in criminal law: the prior-acts and similar-crimes doctrines.[93] Under both of these doctrines, past similar behavior on the part of the defendant can be used as evidence supporting conviction.[94] But these two doctrines, termed the "pattern-of-behavior" doctrines, are distinct from the aggregation discussed above. Whereas the pattern-of-behavior doctrines are based on the probabilistic *dependence* of the offenses attributed to the defendant, the

---

89. For more objections and responses, see Harel & Porat, *supra* note 8, at 291-309.

90. *See* Maya Bar-Hillel, *Probabilistic Analysis in Legal Factfinding*, 56 ACTA PSYCHOLOGICA 267, 268-70, 282-83 (1984) (analyzing the use of probabilities in cases and suggesting a "soft role . . . for probability in the *factfinding* process") (emphasis added). *But see* People v. Collins, 438 P.2d 33, 33, 40-41 (Cal. 1968) (rejecting the use of probabilities in determining guilt on the facts of the case).

91. *Compare* Levmore, *supra* note 8, at 733 n.19 (suggesting that the defense might benefit from a rule of aggregation when it reminds the jury of all the doubts that have been raised and implies that, combined, they create more than a reasonable doubt), *with* Nash, *supra* note 8, at 138-39 (discussing the rule of aggregation in the context of voting by judges in a panel or by jurors and observing that "a criminal defendant cannot be convicted unless a jury unanimously finds each element of the crime charged proven beyond a reasonable doubt" (citation omitted)). Note that the Model Penal Code says that "[n]o person may be convicted of an offense unless each element of such offense is proved beyond a reasonable doubt." MODEL PENAL CODE § 1.12(1) (1962). It thus seems that the Code rules out cross-element aggregation.

92. *See* Nash, *supra* note 8, at 138-39 ("[A] federal jury need not always decide unanimously which of several possible sets of underlying brute facts make up a particular element, say, which of several possible means the defendant used to commit an element of the crime." (quoting Richardson v. United States, 526 U.S. 813, 817 (1999))).

93. *See* FED. R. EVID. 404, 413, 414.

94. *See id.*

aggregation we have discussed is most appropriately (but not only) applied when those offenses are entirely *independent* of one another.

Under the prior-acts doctrine, which was adopted in Rule 404(b) of the Federal Rules of Evidence, the prosecution can bring evidence of other crimes, wrongs, or acts that can be attributed to the defendant to establish "motive, opportunity, intent, preparation, plan, knowledge, identity, absence of mistake, or lack of accident."[95] Such prior acts cannot be used as propensity evidence,[96] however, and courts are required to instruct the jury accordingly.[97] Interestingly, under Rule 404(b), as interpreted by the Supreme Court, even conduct that has been the subject of a prior acquittal can be submitted as evidence by the prosecution in a subsequent trial in order to support conviction.[98]

The similar-crimes doctrine, adopted in Rules 413 and 414 of the Federal Rules of Evidence, applies to sexual assault and child molestation offenses.[99] Under this doctrine, if the defendant is accused of one of these types of offenses, "the court may admit evidence that the defendant committed any other" offense of the same type, and "[t]he evidence may be considered on any matter to which it is relevant."[100]

The superficial similarity between the pattern-of-behavior doctrines and aggregation stems from their shared feature, namely, that all three consider the past behavior of the defendant and affirm that past behavior can influence the

---

95. *Id*. R. 404(b)(2).

96. *Id*. R. 404(b)(1).

97. *See, e.g.*, People v. Quinn, 486 N.W.2d 139, 140 (Mich. Ct. App. 1992) ("Where, however, evidence of a defendant's other wrongful acts has been admitted for the limited purposes allowed under [Michigan Rules of Evidence] 404(b), the prosecutor deprives the defendant of a fair trial in arguing that the jury should consider the evidence as substantive evidence of the defendant's guilt."); *see also* Huddleston v. United States, 485 U.S. 681, 689-92 (1988) (holding that the prosecution can admit evidence of past acts under Rule 404(b) without first establishing by the preponderance of the evidence that the defendant committed those acts).

98. *See, e.g.*, Dowling v. United States, 493 U.S. 342, 348-49 (1990) (holding that testimony tending to prove that the defendant had committed a crime, which had been brought in a prior trial that ended in acquittal, was rightly admitted under Rule 404(b) by the court in a subsequent trial because it established the defendant's identity).

99. *See* FED. R. EVID. 413, 414.

100. *Id*. Under Rule 415 of the Federal Rules of Evidence, this doctrine is also applicable to civil cases involving sexual assault and child molestation. *See* Louis M. Natali, Jr. & R. Stephen Stigall, *"Are You Going to Arraign His Whole Life?": How Sexual Propensity Evidence Violates the Due Process Clause*, 28 LOY. U. CHI. L.J. 1, 29 (1996) ("By requiring the admission of propensity evidence, the rules prevent a fundamentally fair trial, and thus violate due process . . . .").

likelihood of conviction.[101] But, this resemblance notwithstanding, there is a substantial difference between them. The pattern-of-behavior doctrines are rooted in the premise that a person who has committed several offenses in the past could be more likely to have either intended to commit or actually committed the offense of which that person is presently accused. The defendant's past behavior thus modifies the probability of his guilt in the current case. It is the *dependence* between the past offense and the present alleged offense that provides the grounds for conviction. In contrast, cross-claim factual aggregation is based on the axiom that the probability that a person committed at least one of two offenses (*A* or *B*) is greater than the probability that she committed *A* and greater than the probability that she committed *B* (unless there is full dependence between the two offenses). Cross-claim factual aggregation is not based on any dependence between the offenses attributed to the defendant: in our examples, the probability that she committed one offense does not change the probability that she committed another. Rather, only the probability that she committed *an unspecified offense* is affected.[102]

B. Normative Aggregation

Normative aggregation could arise in criminal law in situations analogous to tort law. Thus, like the tort defendant, the criminal defendant may raise two defenses, neither of them sufficient to exonerate the defendant from liability. However, the weight of those two "almost defenses" taken together may be sufficient for acquittal, or at least for mitigation of liability. For example, the defendant may raise both a factual mistake defense and self-defense. Suppose that neither of the two defenses reaches the normative threshold where it applies: the mistake was unreasonable (but nearly reasonable) and the

---

101. As Example III.2 illustrates, sometimes aggregation would lead to acquittal rather than conviction.

102. Note that under the prior-acts and similar-crimes doctrines, the fact that a person committed several similar offenses in the past increases the chances of conviction in the present case. By contrast, under cross-claim factual aggregation, as illustrated by Example III.2, *if it were practical to apply aggregation in different trials*—an issue we discuss at greater length later, *see infra* Subsection V.C.2—the fact that a person was convicted of an offense in the past in one trial would decrease the probability of conviction in a later offense in a subsequent trial. Thus, if in Example III.2 the defendant were convicted of rape in a first trial (because the probability of his guilt was 95%) and then charged with pickpocketing in a subsequent trial, a court applying aggregation would acquit the defendant in the later trial as long as the probability of his guilt of pickpocketing was lower than 100%.

defendant used unreasonable (but nearly reasonable) force.[103] Aggregating the weights of the two "almost defenses" could lead to the defendant's acquittal. Courts, however, do not aggregate defenses, at least not explicitly.

But, even more interestingly, cross-claim normative aggregation could also be made across several criminal charges, when the same defendant committed several "almost offenses," and the aggregation of those "almost offenses" could justify conviction. The next example illustrates such a case:

> *Example III.3. Several Minor Misdeeds.* The defendant is accused of five separate offenses, allegedly committed in different times and places, of interrupting the work of a public official. None of the acts considered separately reaches the point where the behavior is defined as an offense.

Under prevailing law, the defendant would be acquitted of the five charges brought against him.[104] But if all cases were aggregated, the court could reach a different decision. Thus, if the behavior in each case is reprehensible, but not reprehensible enough to justify the application of the criminal law, the cumulative weight of all five cases could be more than enough to justify such an application. We can think of two main reasons why five cases could justify

---

103. In *State v. Thornton*, 730 S.W.2d 309 (Tenn. 1987), the defendant was charged with murder after shooting a stranger whom he discovered in bed with his wife. *Id.* at 309. The defendant raised the affirmative defenses of insanity and self-defense. *Id.* at 312. He lost on both defenses at trial, and was convicted of murder in the first degree, but the appellate court, citing the particular emotional stress of the defendant's encounter with the victim, reduced his conviction to voluntary manslaughter. *Id.* at 312-15. One possible explanation for the appellate court's decision is that the court implicitly engaged in normative aggregation of his two "almost defenses." In *Johnson v. State*, 36 So. 3d 170 (Fla. Dist. Ct. App. 2010), the defendant raised two defense claims: that he was not present at the place where the murder took place, and that the codefendant's fatal beating of the victim was an unforeseen independent act falling outside of the original plan of the crime. *Id.* at 171. The jury denied both defenses and convicted him of second-degree murder, and the appellate court subsequently upheld the trial court's jury instruction and thus the verdict. *Id.* at 171-72. From the facts as presented by the appellate court, it seems possible that one defense (the claim that the defendant was not present when the killing took place) was not established with sufficient factual certainty, while the other defense (that the killing constituted an unforeseen independent act) fell short of the requisite normative threshold of unforeseeability. *Id.* If this reading of the case is right, then a mixed aggregation (factual and normative) might have brought a different result.

104. *See, e.g.*, State v. Stolen, 755 N.W.2d 596, 603 (Neb. 2008) (reversing the accused's conviction on the ground that physical interference is required in order to commit obstruction of government operations). *But cf.* Duncantell v. State, 230 S.W.3d 835, 841-2 (Tex. Ct. App. 2007) (affirming the defendant's conviction for interference with the duties of a public servant because he disregarded the officer's requests to stand back by *repeatedly* entering the crime scene area).

conviction even if one case does not: first, it may be that applying criminal law to one occurrence only is not cost-justified, while with five occurrences it is cost-justified. Second, the recurrence of the same event five times may shed new light on the wrongfulness of the defendant's behavior as a whole and may justify convicting him.

Criminal law has a number of aggregation doctrines that allow courts to normatively aggregate separate offenses (or "almost offenses") so as to create an entirely new offense.

One example is the offense of stalking. Under anti-stalking acts, a single act of harassment does not constitute an offense, but if that behavior occurs several times, then at a certain point it becomes an offense.[105] The New Jersey Criminal Code, for example, defines stalking as "repeatedly maintaining a visual or physical proximity to a person."[106] Thus, the Code arguably formalizes a type of normative aggregation: it instructs courts that one act of harassment is bad, but not bad enough to justify the law's intervention, while several such acts might cumulatively justify a conviction. A second interpretation is that one act of harassment could be accidental, with no malicious motives, and therefore, in order to reduce the risk of false convictions, the law requires more than one act of harassment.[107]

In other cases, a single act is an offense, but if that behavior is repeated several times, this series of offenses can constitute a more severe offense. That could also be regarded as normative aggregation. A typical example is the importation of drugs: if the accused imports drugs once and the quantity is small enough, he would be convicted of the offense of drug possession of the first (lowest) degree. If, however, the accused imports a large enough quantity, then he would be convicted of a higher degree of the offense. The large-quantity condition could be satisfied even if the accused imports drugs several times, each time moving only a small quantity that would qualify

---

105. *See* Heather C. Melton, *Stalking: A Review of the Literature and Direction for the Future*, 25 CRIM. JUST. REV. 246, 247 (2000) ("The term 'stalking' is used to describe the willful, repeated, and malicious following, harassing, or threatening of another person . . . .").

106. N.J. STAT. ANN. § 2C:12-10-a (West 2012). For the New Jersey Superior Court's interpretation of the stalking clause, see *State v. Cardell*, 723 A.2d 111, 114-18 (N.J. Super. Ct. App. Div. 1999).

107. In *State v. Berg*, 213 P.3d 1249, 1250 (Or. 2009) (en banc), the defendant was convicted of tampering with a witness and stalking, "based on allegations that he had repeatedly trespassed on his neighbors' property, engaged in aggressive and offensive conduct toward them, and threatened one of them with various consequences if she 'showed up in court.'" The court focused on one specific event and convicted the defendant. *Id.* at 1250-51. This decision might be interpreted as an implicit aggregation of the numerous misbehaviors of the defendant, when each of them standing alone would not constitute an offense.

independently for the first degree, as long as the total quantity across all occasions is large enough.[108]

Another example of normative aggregation is the doctrine embedded in the Racketeer Influenced and Corrupt Organization Act (RICO).[109] Under RICO, a person who is a member of an enterprise that has committed any two specified crimes within a ten-year period can be charged with racketeering.[110] Thus, the offense of racketeering can be characterized as a result of normative aggregation of two separate offenses (committed by an enterprise) which can underlie a new offense (committed by an individual) of belonging to that criminal enterprise.

The use of normative aggregation to address cooperative criminal activity has cropped up in other settings. Police often have a good idea of who the local mischief-makers are. In big cities, these people often belong to criminal gangs. When rival gangs fight over turf and cause disorder, the police have reason to suspect that many members of each gang are involved but will not have sufficient proof to convict anyone of the offense other than those who participate directly. RICO tried to address this problem, but proving the two separate crimes that are necessary for applying RICO can also be too high a hurdle. Other laws, such as Chicago's gang loitering law,[111] attempted to address this problem indirectly by permitting the police to disperse groups of

---

108. In *United States v. Shonubi*, 802 F. Supp. 859, 860 (E.D.N.Y. 1992), *rev'd*, 998 F.2d 84 (2d Cir. 1993), the defendant, arriving from Nigeria, was arrested at JFK International Airport with 427.4 grams of heroin. He was charged with importing heroin and possessing heroin with the intent to distribute it. The court found that in addition to the last occasion when the defendant was arrested, he had made seven other trips to Nigeria. The court concluded that those seven trips had been made for the purpose of importing heroin, and therefore multiplied the quantity of 427.4 grams imported on the time of arrest by eight. *Id.* at 860-61. The defendant was convicted and the court, applying the drug quantity table of the sentencing guidelines, classified the case as falling under "level 34" of the table, which relates to a drug quantity between 3 and 10 kilograms, instead of "level 28" of the table, which relates to a drug quantity between 400 and 700 grams. United States v. Shonubi, 103 F.3d 1085, 1087 (2d Cir. 1997). That classification of the offense allowed the court to impose on the defendant a much harsher sentence than if the case were classified under "level 28" of the table. *Id.* The court's decision was later vacated and remanded twice by the Second Circuit, first in *United States v. Shonubi*, 998 F.2d 84 (2d Cir. 1993), and later in 103 F.3d 1085 (2d Cir. 1997), after the district court had reinstated the same verdict on remand.

109. 18 U.S.C. §§ 1961-1968 (2006).

110. *Id.* § 1961(5).

111. This law was struck down by the Supreme Court in *City of Chicago v. Morales*, 527 U.S. 41 (1999). Chicago passed a new anti-gang ordinance after the Supreme Court decided *Morales*. *See* Dirk Johnson, *Chicago Council Tries Anew with Anti-Gang Ordinance*, N.Y. TIMES, Feb. 22, 2000, http://www.nytimes.com/2000/02/22/us/chicago-council-tries-anew -with-anti-gang-ordinance.html.

people if a known gang member was present, and arrest anyone who failed to comply with orders to disperse.[112] We suspect that this law enabled police to, in effect, aggregate claims against known mischief-makers—people who had committed minor acts of wrongdoing (normative aggregation) or were reasonably suspected of having been involved in serious offenses (factual aggregation).

C. Cross-Person Aggregation

Cross-person aggregation is largely absent in criminal law, probably because it could infringe on the accused's constitutional rights. Thus, if there are several defendants accused of committing several crimes, none of them will be convicted even if statistically each of them probably committed some of the crimes. A market share liability approach, applied by some jurisdictions to tort cases (mainly in the DES cases[113]), is unlikely to be considered suitable for criminal cases. In criminal trials, the prosecution must prove the defendant's guilt beyond a reasonable doubt, and statistical evidence cannot be the main evidence for conviction.[114]

However, RICO can be understood as a form of cross-person aggregation. RICO is frequently used to target racketeering offenses in which a large number of people are victims of minor offenses like drug crimes and prostitution. Prosecutors enjoy considerable discretion, and they might choose not to prosecute the crimes individually because the harm to each victim is relatively minor. RICO enables them to aggregate the offenses, perhaps on the theory that the individually minor harms should be considered significant when aggregated across victims. This type of cross-person aggregation might therefore be considered normative, but there is also a factual version. Suppose that we cannot identify which of a number of gang members committed certain crimes, but we can convict all of them of belonging to a gang involved in a criminal enterprise. As a result, some gang members may be, in effect, convicted of the crimes committed by other gang members. Here, RICO permits factual aggregation across persons: we cannot connect any particular member to any particular victims with confidence, but we can be confident that all the gang members committed a crime against at least some of the victims.

---

112. *Morales*, 527 U.S. at 47.

113. *See supra* notes 54-57 and accompanying text.

114. *See* STEIN, *supra* note 24, at 183-85 (explaining why statistical evidence, standing alone, cannot be the basis for conviction).

This logic is most clearly visible in the doctrine of joint criminal enterprise liability in international criminal law. The doctrine was recognized by the International Tribunal for the Former Yugoslavia in the *Tadic* case.[115] The Appeals Chamber held that a defendant could be convicted for taking part in a joint criminal enterprise that foreseeably resulted in the killing of victims, even though the original purpose of the enterprise was to rid the region of the non-Serbian population, not specifically to kill, and the defendant could not be directly tied to the killing.[116] The foreseeability requirement is inherently probabilistic: if a person joins a group with an agenda of causing mayhem, that person's expected liability increases with the number of expected victims.

Joint criminal enterprise liability has been criticized by criminal law scholars and some international law scholars who believe that it erodes the procedural protections of defendants. They argue that sympathy for the victims of mass atrocity has caused governments and international lawyers to endorse international criminal law doctrines that are unfair to defendants.[117] From another perspective, however, the use of aggregation rules for mass atrocities makes good sense. Normative aggregation may be justified because relatively weak forms of complicity that are not blameworthy in normal times may be considered blameworthy in times of mass atrocity even if one does not go so far as to endorse collective responsibility. Factual aggregation may also be justified because the large number of victims may in itself raise the probability that a defendant was criminally involved, as in the *Tadic* case.

IV. PUBLIC LAW

A. Factual Aggregation

In public law, problems of cross-claim factual aggregation arise in numerous settings. Consider the following example:

---

115. Prosecutor v. Tadic, Case No. IT-94-1-A, Judgment (Int'l Crim. Trib. for the Former Yugoslavia July 15, 1999).

116. *Id.* ¶¶ 230-232. For a lucid discussion of the *Tadic* case and joint criminal enterprise liability, see Allison Marston Danner & Jenny S. Martinez, *Guilty Associations: Joint Criminal Enterprise, Command Responsibility, and the Development of International Criminal Law*, 93 CALIF. L. REV. 75, 104-07 (2005).

117. *See, e.g.*, Danner & Martinez, *supra* note 116; George P. Fletcher & Jens David Ohlin, *Reclaiming Fundamental Principles of Criminal Law in the Darfur Case*, 3 J. INT'L CRIM. JUST. 539 (2005).

Example IV.1. Targeted Killing: Alternative Claims. The President seeks to use military force to kill a suspected terrorist in Pakistan. However, there is uncertainty both about whether the suspect is planning an attack on U.S. targets and about whether the person is a member of Al Qaeda. Assume that killing the person is lawful if either condition is valid. Suppose that the probability of each independently is 40% and that the law requires a probability of more than 50% to justify the killing.

We suspect that the President's lawyers would advise him that he cannot order the killing of the target. But if aggregation were accepted, the probability that the killing would be lawful would be 64%, and thus the correct legal advice would be the opposite.

Now consider a case of cross-element factual aggregation.

Example IV.2. Targeted Killing: Cumulative Elements. Same as Example IV.1, except suppose that the law provides that the killing is lawful only if the target is a non-American *and* the target is planning an attack. The probability that the target is a non-American is 60%, and the probability that the target is planning an attack is 60%.

We believe that the President's lawyers would advise him that he cannot order the killing of the target because the probability that the target is not an American *and* is planning an attack is only 36% (at least if the probabilities are independent—and they may not be). If we are right, then factual aggregation of several claims is applied inconsistently—barred in the first case, required in the second case. Our minimal suggestion is that factual aggregation of claims should be used consistently—if the President cannot order the killing in the second case, then he must be permitted to do so in the first case.[118]

B. Normative Aggregation

As in the other legal settings we have examined, we can imagine cases where cross-claim normative aggregation could occur in public law.

Example IV.3. Targeted Killing: Two "Almost Claims." The President seeks to use military force to kill an American citizen who is alleged to be associated with Al Qaeda and who lives in Sana'a, the capital of

---

118. Note that Example IV.1 involves cross-claim aggregation and Example IV.2 involves cross-element aggregation. Although in other contexts these types of aggregation might properly be treated differently, we see no justification for different treatment in these two examples.

Yemen. The President claims two sources of authority: a statute that gives him the authority to use military force against Al Qaeda, and his constitutional power to use military force abroad to protect American interests. Each claim is at best controversial—many contest that the statute authorizes action outside a conventional battlefield, and that the President can use his constitutional powers to kill an American citizen.

We suspect that most commentators believe that the President could use military force in this case only if at least one of his claims standing alone could be established. Consequently, in Example IV.3, the President would not be permitted to order a targeted killing. Under the aggregation approach, one might reason differently. The President has two "almost claims": that under the statute he can use force against a terrorist in foreign territory beyond the control of domestic law enforcement authorities—even if not on the "battlefield," strictly speaking—and that under the Constitution he can use force against enemies abroad to protect American interests, even if the enemy is an American citizen. If we aggregate the normative weight of these "almost claims," then the President arguably possesses the authority to order a targeted killing in Example IV.3.

This argument might seem fanciful, but it is fairly common in constitutional adjudication involving the authority of the executive. In *Dames & Moore v. Regan*,[119] for example, the Supreme Court affirmed the President's authority to suspend American claims against Iran based on an aggregation of statutory and constitutional powers:

Although we have declined to conclude that the [International Emergency Economic Powers Act] IEEPA or the Hostage Act directly authorizes the President's suspension of claims for the reasons noted, we cannot ignore the general tenor of Congress' legislation in this area in trying to determine whether the President is acting alone or at least with the acceptance of Congress. As we have noted, Congress cannot anticipate and legislate with regard to every possible action the President may find it necessary to take or every possible situation in which he might act. Such failure of Congress specifically to delegate authority does not, "especially . . . in the areas of foreign policy and national security," imply "congressional disapproval" of action taken by the Executive. On the contrary, the enactment of legislation closely related to the question of the President's authority in a particular case which evinces legislative intent to accord the President broad discretion

---

119. 453 U.S. 654 (1981).

may be considered to "invite" "measures on independent presidential responsibility [. . . .]"[120]

The two statutes, the IEEPA and the Hostage Act, did not independently authorize the President's action, but each almost did. That, along with the President's constitutional authority in the area (which also does not by itself authorize the action, but almost does), in the aggregate provided the President with the claimed authority.

Another line of cases endorses cross-claim normative aggregation where plaintiffs allege violations of the right to free exercise of religion.

> *Example IV.4. Free Exercise of Religion and Free Speech.* A church challenges a zoning ordinance that provides that only industrial structures may be built in an area of a city. The church argues that the zoning ordinance violates both its constitutional right to free exercise of religion and its constitutional right to free speech. Taken separately, the claims would fail. The zoning ordinance is a valid neutral law that does not discriminate against religious organizations, and it does not put an unreasonable burden on speech.

In *Employment Division v. Smith*,[121] the Supreme Court upheld a statute denying unemployment benefits to a person who had illegally used peyote in a religious ritual. Distinguishing (on controversial grounds) an earlier precedent that held that laws that burden the free exercise of religion are subject to strict scrutiny,[122] the Court held that any "neutral" and "generally applicable" law survives constitutional challenge under the Free Exercise Clause even if it incidentally burdens religious practice.[123] However, the Court also recognized a "hybrid" exception. Where a plaintiff can show that a neutral law burdens both religious practice and another constitutionally protected activity, the law is subject to strict scrutiny, and therefore will be struck down unless the government can show a compelling state interest.[124]

---

120. *Id.* at 678 (citations omitted).

121. 494 U.S. 872 (1990).

122. *Id.* at 882-85 (distinguishing Sherbert v. Verner, 374 U.S. 398 (1963)).

123. *Id.* at 881-82.

124. *Id.* The Court did not apply the hybrid exception to the plaintiff's claim, presumably because the plaintiff alleged that only one constitutional norm was violated. *Id.* The Court used the hybrid exception to distinguish, among other cases, *Wisconsin v. Yoder*, 406 U.S. 205 (1972), where the Court struck down a neutral law because of the burden it imposed on religious association. *Smith*, 494 U.S. at 881. The Court's reasoning has been harshly criticized by numerous commentators because, among other reasons, *Yoder* itself did not

The hybrid rights exception fits our definition of cross-claim normative aggregation. In the words of one scholar, "[A] less than sufficient free exercise claim, plus a less than sufficient claim arising under a different part of the Constitution, together trigger the compelling interest test."[125] More formally, consider a claim that a statute violates two provisions of the Constitution, *X* and *Y*, where the plaintiff can show that the statute does not serve a compelling state interest under the strict scrutiny test. Although the statute does not violate *X* or *Y* individually, it does violate them jointly, and thus would be struck down.

Applying the hybrid exception to Example IV.4 could result in accepting the church's claim and overriding the zoning ordinance. The lower courts have heard numerous hybrid cases similar to our example.[126] Churches have frequently challenged zoning ordinances on the grounds that the ordinances violate the Free Exercise Clause and the Free Speech Clause (or the Equal Protection Clause). Each claim is individually weak: zoning ordinances are usually neutral and generally applicable—for example, an ordinance might permit only industrial buildings in an area where people want to build a church—and so do not violate the Free Exercise Clause by themselves. And zoning ordinances are rarely held to violate the Free Speech Clause because the ability to speak to an audience does not depend on having a building in a particular area. Yet, under the hybrid rights approach, a church could in theory prevail as long as each individually losing claim is "colorable" or exceeds some threshold of plausibility.[127]

---

mention hybrid rights—the outcome was based solely on the Free Exercise Clause. *See, e.g.*, Michael W. McConnell, *Free Exercise Revisionism and the* Smith *Decision*, 57 U. CHI. L. REV. 1109, 1121-22 (1990).

125. Richard F. Duncan, *Who Wants To Stop the Church: Homosexual Rights Legislation, Public Policy, and Religious Freedom*, 69 NOTRE DAME L. REV. 393, 430-31 (1994).

126. For an excellent but dated survey, see William L. Esser IV, Note, *Religious Hybrids in the Lower Courts: Free Exercise Plus or Constitutional Smoke Screen?*, 74 NOTRE DAME L. REV. 211 (1998). Esser, however, concludes that the hybrid claims prevail only when the constitutional claim with which a Free Exercise claim is paired would independently prevail. *Id.* at 242-43. The literature on hybrid rights is quite negative. *See, e.g.*, Steven H. Aden & Lee J. Strang, *When a "Rule" Doesn't Rule: The Failure of the* Oregon Employment Division v. Smith *"Hybrid Rights Exception,"* 108 PENN ST. L. REV. 573 (2003); McConnell, *supra* note 124; Kyle Still, Smith*'s Hybrid Rights Doctrine and the* Pierce *Right: An Unintelligent Design*, 85 N.C. L. REV. 385 (2006). These authors argue that the doctrine makes no sense on its own terms and has sown confusion among the lower courts. The *Smith*-related case law has been complicated by the effect of the Religious Freedom Restoration Act, 42 U.S.C. §§ 2000bb to 2000bb-4 (2006), which we ignore because of our focus on constitutional issues.

127. *See, e.g.*, Swanson *ex rel.* Swanson v. Guthrie Indep. Sch. Dist. No. I-L, 135 F.3d 694, 700 (10th Cir. 1998) (requiring that a showing of an infringement of constitutional rights be

Another group of examples involves challenges to laws that allegedly infringe on the parental right to educate one's children. The right to educate one's children is a constitutional right, but a weak one, and parents usually fail when they challenge truancy laws and schools' educational policies on the basis of this right. But when parents claim that educational laws infringe on both their parental rights *and* their free exercise rights, even though the laws are neutral and generally applicable, they make out a hybrid claim and may obtain relief.[128]

Outside of free exercise, it is difficult to find clear examples of the recognition of hybrid rights, but in a number of cases the doctrinal logic suggests such a theory. In *Boy Scouts of America v. Dale*,[129] the Court struck down a state law that forced the Boy Scouts to admit a gay counselor in violation of that organization's bylaws. The Court held that the statute violated the Boy Scouts members' rights to "expressive association," a right which might be taken as a hybrid of the right to free speech and the right to association.[130] In *Griswold v. Connecticut*,[131] the Court struck down a statute that prohibited the sale of contraceptives. In a much-criticized opinion, the Court held that the statute was unconstitutional because it violated a right to privacy derived from the "emanations" of a number of different rights in the Constitution, including rights under the First, Third, and Fourth Amendments.[132] Because the Court did not hold that the statute violated any of these rights individually, the implication is that the statute was unconstitutional only because it violated those rights jointly, although the opinion certainly does not make this argument explicitly. Arguably, in *Roberts v. United States Jaycees*,[133] the Court derived a right of intimate association (such as noninterference in family life) from the right to association in the First

---

"colorable"); Cornerstone Bible Church v. City of Hastings, 948 F.2d 464, 473 (8th Cir. 1991) (instructing the district court, on remand, to construe a church's challenge to a zoning ordinance as a hybrid claim entitled to strict scrutiny).

128. *See, e.g.*, People v. DeJonge, 501 N.W.2d 127 (Mich. 1993) (holding in favor of a hybrid claim that a law that required teachers to be certified, and thus interfered with homeschooling, violated free exercise and parental control rights).

129. 530 U.S. 640 (2000).

130. *Id.* at 644. A somewhat similar case is *Hudson v. Craven*, 403 F.3d 691 (9th Cir. 2005), where the Court held that a teacher's claim that she was discharged for attending an anti-WTO rally with a group of students was a "hybrid speech/association claim," which raised questions as to whether the constitutional standard based only on violations of speech rights should be applied. *Id.* at 696.

131. 381 U.S. 479 (1965).

132. *Id.* at 484.

133. 468 U.S. 609, 618-19 (1984).

Amendment and the right to due process in the Fourteenth Amendment.[134] Finally, it has been argued that several recent Supreme Court cases are best understood as reflecting a hybrid claim involving due process and equal protection rights.[135] In *Lawrence v. Texas*,[136] for example, the Court resisted the equal protection argument that homosexuals form a suspect class, and the due process argument that a law prohibiting homosexual sodomy is substantively irrational—but, combining concerns reflected in both clauses, concluded that the statute was unconstitutional.[137]

Thus, we can identify two sorts of hybrid rights cases. The first is where the constitutional claims are treated as separate, but a remedy is granted if each claim is "colorable" or crosses some other threshold. The second is where the courts develop the doctrine, creating a new right by combining two or more recognized rights. *Smith* illustrates the first approach: the Court refrained from recognizing a new right to, say, "parental-religious control." The right to privacy recognized in *Griswold* illustrates the creation of a new right on the basis of two or more recognized rights. This doctrinal evolution is analogous to the way the unconscionability doctrine was developed so as to aggregate previously recognized claims.

At the same time, it is important to emphasize that outside these settings courts rarely respond sympathetically to hybrid claims. In *Wilkie v. Robbins*,[138] for example, the owner of a ranch claimed that officials from the Bureau of Land Management engaged in a campaign of harassment over a number of years, including trespasses and malicious prosecutions, in an effort to compel him to grant an easement to the U.S. government. The rancher brought two Fifth Amendment claims, describing the Bureau's campaign as both a taking and an illegal form of retaliation.[139] Of these, the Court appears to have considered only his retaliation claim. While the Court acknowledged serious government misconduct, it ultimately declined to allow a *Bivens* action for retaliation under the Fifth Amendment.[140] What is notable about this case is that the Court evidently believed that the Bureau had acted wrongfully, but did not consider the possibility that even if the Bureau did not violate two separate

---

134. *See* Marcum v. Catron, 70 F. Supp. 2d 728, 733-34 (E.D. Ky. 1999) (concluding that *Roberts* defined intimate association as a "hybrid right").

135. *See* Kenji Yoshino, *The New Equal Protection*, 124 HARV. L. REV. 747, 750 (2011).

136. 539 U.S. 558 (2003).

137. Yoshino, *supra* note 135, at 778-79.

138. 551 U.S. 537, 541-43 (2007).

139. *Id.* at 547-48.

140. *Id.* at 561-62.

Fifth Amendment rights individually, it did violate the two of them taken together. This approach is the norm – the hybrid rights cases are the exceptional cases. Plaintiffs frequently argue constitutional rights violations in the alternative and, outside the cases we discuss above, courts rarely address the possibility that individually weak claims may be jointly strong.

For another example, consider *United States v. Sanders*,[141] a case in which a defendant was sentenced to a term of thirty-seven months for committing a crime, was released at the end of his sentence, and then was sent back to prison four years later after an appellate court (following substantial delays) determined that his sentence should have been 180 months.[142] The defendant argued that re-imprisonment after such a delay violated his rights to substantive and procedural due process.[143] Other courts had held that a defendant who is sent back to prison as a result of an administrative error could have a substantive due process claim based on the fact that he or she had developed an expectation as to the finality of the sentence and that this expectation was unfairly disappointed. By contrast, the Sixth Circuit held that Sanders's sentence had been appealed by both sides, so Sanders had no reason to believe that his sentence was final.[144] The court also rejected Sanders's procedural due process claim, noting that although the four-year delay was severe, Sanders could not show that it resulted from bad faith or an attempt by the government to gain a tactical advantage.[145] The substantive due process and procedural due process claims were both colorable, but individually they were too weak to warrant relief.

The dissent argued that Sanders should be released, based on an analysis that, in the majority's words, "seems to conflate the procedural due process factors . . . with the substantive due process right . . . to create a sort of hybrid right not to be returned to prison."[146] The dissent, in essence, argued that even if the substantive due process violation was not as serious as in other cases (Sanders's expectation about his sentence should not have been as "crystallized" as in a clerical error case since the sentence was on appeal), and the procedural due process claim was not as serious as in other cases (the delay was significant but not caused by bad faith), the violations jointly considered

---

141. 452 F.3d 572 (6th Cir. 2006).

142. *Id.* at 574.

143. *Id.* at 576.

144. *Id.* at 576-77.

145. *Id.* at 581-82.

146. *Id.* at 583.

entitled Sanders to relief.[147] Like the Supreme Court in the *Roberts* case, the dissent sought to assert a new hybrid right that was based on two recognized constitutional rights, each of which was independently too weak to justify a constitutional remedy.

C. Cross-Person Aggregation

As we saw in Part I, tort law permits aggregation only on occasion, and otherwise falls well short of what aggregation would require. Recall, for instance, Example I.8, which involved a mass tort with indeterminate victims. The defendant's factory pollutes, creating a statistical likelihood of harming an additional 25 people per year, but none of those people can be identified. Tort law does not permit cross-person aggregation, so a remedy is not available in this case. Tort law also does not usually permit cross-claim factual aggregation and normative aggregation. Because of these limitations, Congress and state legislatures have enacted numerous statutes that regulate behavior that otherwise slips through tort law. This is an important domain of public law. Indeed, this type of regulation is ubiquitous—consider speed limits, for example, which protect unidentified future victims by regulating ex ante—and the proposition that public law overcomes the anti-aggregation bias in private law by permitting cross-person aggregation is understood in the literature, even if not put in those terms.[148] But it is worth dwelling on this point, for it shows clearly that the anti-aggregation bias in private law is (at least with respect to cross-person aggregation) not based on any fundamental moral commitments.

For example, the Environmental Protection Agency will typically identify sources of pollution such as factories, and conduct studies that determine whether the pollution emitted by those factories causes harm. The Agency can rarely identify particular people who have been harmed because of the difficulty of untangling other causal factors. But the Agency can use statistical techniques to determine the difference between the number of cases of, say, lung cancer in the population exposed to the pollution as well as to background factors, and the number of cases of lung cancer in a population exposed to the background factors alone. If the difference is large enough, the Agency will issue regulations requiring the factories to reduce their pollution. No victim

---

147. *United States v. Sanders*, 452 F.3d 572, 589-93 (6th Cir. 2006).

148. *See, e.g.*, Steven Shavell, *Liability for Harm Versus Regulation of Safety*, 13 J. LEGAL STUD. 357, 363 (1984) (noting that regulation may be superior to the tort system when victims cannot identify wrongdoers).

receives a remedy, but in the future there will be fewer victims. This is a clear example of cross-person aggregation.[149]

Public law also aggregates across persons. Courts have relaxed standing doctrine so as to allow governments, agencies, and even individuals to bring tort actions on behalf of large groups of people who may have been harmed only in a statistical sense. Among many such cases,[150] a prominent example involves religious displays on public property, which are forbidden under the First Amendment if they are sectarian.[151] One way of thinking about religious displays is that they inflict a nonphysical, and hence difficult-to-prove but nonetheless important, psychological harm on a group of people outside the religion that the display celebrates. If only one religious dissenter saw a display, she would be unable to prove that she was harmed, but if many religious dissenters saw the display—as the "public" nature of the display implies—it is statistically likely that at least one of them was harmed. We believe this statistical likelihood, itself rooted in cross-person aggregation, is the reason why the law gives even a single individual standing to bring a claim. Such an individual in effect sues on behalf of the group, given that a publicly displayed religious symbol by definition reaches a multitude of observers, even if in imaginable circumstances the display is so remotely located that only a few people will ever see it. This approach can be contrasted to the law's reluctance to give remedies to tort victims who claim emotional but not physical harm.[152] The difference is that religious displays by their nature are observed by large groups of people so that small likelihoods of harm can be aggregated, whereas torts that cause emotional harms generally involve only single victims.[153]

---

149. More specifically, it is a kind of cross-person factual aggregation; the victims are (on average) harmed but cannot prove their harm above the requisite probability threshold individually, though they can collectively.

150. *See, e.g.*, Massachusetts v. EPA, 549 U.S. 497 (2007) (holding that the state has standing to challenge the EPA's refusal to issue greenhouse gas regulations where the harm caused by climate change does not result in a particular injury to identifiable people); Georgia v. Tenn. Copper Co., 206 U.S. 230 (1907) (holding that the state has standing to bring a tort claim against a polluter on behalf of its citizens).

151. *See* Cnty. of Allegheny v. ACLU, 492 U.S. 573 (1989); Lemon v. Kurtzman, 403 U.S. 602 (1971).

152. *See* EPSTEIN, *supra* note 41, at 274-75 (noting that recovery for emotional harm without physical impact is generally denied everywhere today).

153. *But see supra* notes 58-61 and accompanying text.

V. EXPLANATIONS AND PROPOSALS

A. The Arbitrariness of Legal Boundaries

The best explanation for courts' inconsistent approaches to these aggregation puzzles is that the division of the legal system into bodies of law, and then those bodies of law into separate claims, and then again those claims into elements, brings costs as well as benefits. Courts respond to those costs by aggregating under certain circumstances, but because they respond in a cautious, ad hoc way, fearful of sacrificing the benefits of disaggregation, the law as a whole contains many inconsistencies.

To understand this problem, we start with factual aggregation and return to our first example from tort law. The general purposes of tort law are twofold: to optimally deter people from imposing externalities on each other and, in some cases, to compensate people who have suffered from those externalities.[154] An ideal decisionmaker who faced no decision costs could be given a simple instruction, such as "maximize social welfare" or "minimize social costs." Such a decisionmaker would aggregate harms and probabilities to fulfill the purpose of tort law in all circumstances.[155] Take, for example, the defendant who acted negligently with 40% probability while driving a vehicle that was inherently dangerous with 40% probability (Example I.1: The Inherently Dangerous Vehicle). In a world without decision costs, aggregation would help to efficiently deter the defendant. If a person is considering whether to engage in these actions, and knows that she will not be held liable because her behavior falls between the cracks of two claims that will not be aggregated, she will engage in those actions, even though in an expected sense they will cause harm which is higher than their benefit. By holding such a person liable, aggregation deters socially costly behavior. Now take, as a second example, the defendant hospital whose doctors negligently caused two separate injuries, each with 60% probability (a variation of Example I.3: Injury in the Hospital: Two Events, Two Injuries). Without aggregation, the defendant's expected liability would be excessive and overdeterrence might result.[156] With aggregation, however, a decisionmaker would hold the defendant liable for

---

154. Under a negligence rule, such compensation would be appropriate only if the externalities were inefficiently (or unreasonably) imposed.

155. But not in the very rare cases when aggregation may create over- or underdeterrence. *See supra* note 25.

156. *See supra* notes 20-25 and accompanying text.

only one injury, because the probability that the defendant negligently caused the two injuries is only 36%.

But, as the literature on rules versus standards makes clear,[157] simply telling decisionmakers to maximize social welfare—a pure standard—would not work very well in a world in which decision costs are often high. The preponderance-of-the-evidence rule, applied with no aggregation, greatly simplifies decisionmaking; if courts were required to make point estimates and combine them, then their job would be more difficult. It would also be more difficult for parties to predict the legal consequences of their behavior.[158] Thus, the structure of our current legal system reduces decision costs, but in doing so, sacrifices some of the efficiency gains that result from factual aggregation.

We turn now to normative aggregation. Take Example II.2 (Non-Material Breach and Misleading Conduct), where the defendant engages in misleading conduct in order to secure the plaintiff's consent to a contract and subsequently engages in a breach that falls just short of material. A court would not allow the plaintiff to rescind the contract because each claim is considered separately. Aggregation would create a more complicated rule for both the parties to the contract and the courts, and thus increase decision costs. Rather than determine (1) whether the misleading conduct crosses a normative threshold, and (2) whether the breach crosses a normative threshold, the court would be required to determine (1) whether the misleading conduct crosses a normative threshold, (2) whether the breach crosses a normative threshold, *and* (3) whether the *combined* actions cross a normative threshold. This more complicated test may well create an unacceptable level of difficulty and uncertainty. But nonaggregation also entails costs: it permits a defendant to escape liability for two actions that are jointly, by assumption, inefficient or unjust.

This argument explains why one can more easily find examples of within-element factual aggregation than cross-claim factual aggregation and normative aggregation, and why courts are more likely to aggregate claims of a similar nature than claims of a different nature. Consider, for example, a case where a plaintiff can show that the defendant committed a tort against him with 40% probability and a breach of contract against him with 40% probability. Combining the breach and tort claims raises possible complexities and unintended consequences: for example, if there are different statutes of

---

157. *See, e.g.*, Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557 (1993).

158. For the difficulties of factual aggregation, see Levmore, *supra* note 8, at 726-33. Most of Levmore's examples come from tort law and relate to what we call cross-element aggregation.

limitations for contract claims and tort claims, which statute should be used when the claims are combined? And if the suit is for damages, there could be different damages rules for breach of contract and tort, so a question arises as to which damages rule should be used when the claims are combined.[159] By contrast, when the claims are of the same nature, no such conflicts arise.

Should the law aggregate more than it does? Our minimal suggestion is that the doctrine is currently inconsistent, and that the courts would do well to bring consistency to this area. If normative aggregation in the hybrid rights doctrine is good policy, for example, then normative aggregation should be good policy in other areas of public law, and in private law, too, *mutatis mutandis*. If the lack of normative aggregation in many areas of the law is good policy, then normative aggregation should be purged from Religion Clause jurisprudence and other areas of the law where it appears. The same point can be made about the other forms of aggregation. To be sure, there may be relevant policy differences that explain the apparently inconsistent use of aggregation in different areas of the law, but we have not been able to identify them.

It is also possible to take a more aggressive view that courts should aggregate whenever decision costs are low. When decision costs are low, factual aggregation will generally improve the accuracy of adjudication while not changing substantive law. Normative aggregation should improve substantive law, in the sense of vindicating values and policy choices that are already found in the law, but which defendants can violate if claims are not aggregated. However, it may be difficult for courts to determine when their own decision costs are low or not, and as we note elsewhere, in some cases the benefits from aggregation are minimal,[160] so the gains from pursuing aggregation more broadly may sometimes be outweighed by the costs.

B. Other Explanations for Failures To Aggregate and Possible Objections to Aggregation

There could be other explanations for courts' reluctance to aggregate. Those explanations could also be grounds for objections to aggregation. Most of the explanations—or objections—relate to one type of aggregation but not

---

159. Note that these questions are much less acute when the plaintiff is able to establish both tort and contract claims separately. In that case, the plaintiff would generally be entitled to the remedy that is more favorable to him.

160. See *supra* p. 15 and the discussion in the text accompanying note 66.

others, or to aggregation in one field of the law but not others. Some of the explanations — or objections — are efficiency-related but others are not.

1. Corrective Justice

Under the principles of corrective justice, the defendant should rectify the injustice he inflicted upon the plaintiff through his wrongdoing by compensating her for the harm done. Theorists of corrective justice maintain that it is crucial that the defendant rectifies the injustice done to the plaintiff and not to an unaffected third party.[161] Moreover, under corrective justice principles, the determination of liability should rest upon the relationship between the defendant and plaintiff as doer and sufferer, and anything outside that relationship should be ignored.[162]

How might corrective justice theorists respond to aggregation in private law (mainly in tort law)? We expect corrective justice theorists to oppose cross-person aggregation, since such aggregation would require taking into account wrongs committed toward third parties while determining the remedies available to the plaintiff against the defendant. Thus, in Example I.8 (Mass Torts: Indeterminate Plaintiffs), a factory wrongfully created radiation, and while it can be established that 25 out of 125 people suffered harm due to the radiation, it is impossible for each plaintiff to establish, by the preponderance of the evidence, that her harm is the result of the radiation. Corrective justice theorists would maintain that all suits should be dismissed since in each and every case it is more probable than not that the defendant did not injure the plaintiff. The mere fact that there are many plaintiffs, and that 25 of them probably suffered harm as a result of the defendant's wrongdoing, should be considered under corrective justice as irrelevant to the determination of liability.[163] Market share liability, however, could be reconciled with at least some versions of corrective justice.[164]

---

161. *See* JULES L. COLEMAN, RISK AND WRONGS 380-85 (2002) (discussing the focus of corrective justice on the relationship between the injurer and the victim); ERNEST J. WEINRIB, THE IDEA OF PRIVATE LAW 63-66 (1995) (same).

162. This is an implication of the correlativity requirement, under which liability should be imposed for harms which are the materialization of the risks that defined the injurer's conduct as negligent. *See* WEINRIB, *supra* note 161, at 159 ("The consequences for which the defendant is liable are restricted to those within the risks that render the act wrongful in the first place.").

163. *See id.* at 63-66 (explaining that corrective justice focuses only on the relationship between the injurer and the victim, and implying that those two should be identified); *see also* Stephen R. Perry, *Risk, Harm, and Responsibility*, *in* PHILOSOPHICAL FOUNDATIONS OF TORT

We see no reason for corrective justice theorists to oppose cross-claim factual aggregation or normative aggregation if it relates to one specific event. The harder case is when there are two separate events occurring between the defendant and the plaintiff and neither of them can be established by the preponderance of the evidence to justify liability, but it can be established by the preponderance of the evidence that at least one of them justifies liability. We suspect that corrective justice theorists would oppose such aggregation, arguing that each event should be considered separately and in isolation from the other.[165] Thus, in Example I.3 (Injury in the Hospital: Two Events, Two Injuries), a patient suffered two distinct harms, each of which might have been caused by a different doctor. The hospital is vicariously liable for both doctors' wrongdoing. While the patient cannot establish the liability of either doctor independently, he can establish that at least one of the harms was caused by a doctor's wrongdoing. Corrective justice theorists would probably argue that each event should be considered separately: the determination of liability should be done per event, and not across events (we suspect that corrective justice theorists would persist in that view even if, in our example, the two harms might have been the result of the same doctor's wrongdoing[166]). In any event, while it is possible that corrective justice intuitions may account for limits on some types of aggregation, they cannot explain the bias against aggregation in most cases.

2. Incommensurability

In law and economics it is assumed that all potential outcomes are commensurable and comparable to one another. But there are also different

---

LAW 321, 330-39 (David G. Owen ed., 1995) (arguing against liability for lost chances of recovery in the absence of reliance or lost opportunities for alternative treatment).

164. Corrective justice will support probabilistic recovery when three cumulative conditions are met: (1) the wrongdoers pay for the harm caused by their wrongdoings; (2) the victims are compensated for the harm wrongfully caused to them; and (3) the wrongdoers make payments to or participate in the mechanism that facilitates the compensation of their victims. These three conditions are satisfied in the DES cases, as well as in other cases of recurring wrongs when: a group of wrongdoers inflicts harms numerous times on a group of victims; the harm caused by each wrongdoer and the harm caused to each victim is verifiable; but it is impossible for each victim to prove the identity of the specific wrongdoer, from the group of wrongdoers, who caused her harm. *See* PORAT & STEIN, *supra* note 29, at 132-33.

165. *See* WEINRIB, *supra* note 161, at 75 ("Corrective justice involves the intrinsic unity of the doer and sufferer of the same harm.").

166. Corrective justice might support liability, however, if the two events can be reasonably understood as one event occurring in two stages.

views that consider certain outcomes incommensurable.[167] A possible explanation for courts' reluctance to engage in some aggregations but not in others is their refusal to evaluate claims of different natures according to one common scale.

Take Examples II.1 (Either Material Breach or Fraudulent Misrepresentation) and II.2 (Non-Material Breach and Misleading Conduct). Both examples deal with a case where the plaintiff argues that he was entitled to rescind a contract because the defendant engaged in fraudulent misrepresentation in order to secure the plaintiff's consent to the contract and subsequently materially breached the contract. In Example II.1, there is a 40% probability that each of the two claims holds, while in Example II.2, although the facts are not disputed, neither the misrepresentation nor the breach, standing alone, is severe enough to justify the rescission of the contract. In order to aggregate the two claims in both examples, it seems that courts need a common scale to measure misrepresentation on the one hand and breach on the other hand. Finding such a scale is impossible – or so the commensurability objection would be.

The commensurability objection would probably be more applicable to Example II.2 than to Example II.1. In Example II.1, the court just needs to estimate the probability that the defendant behaved in a way that warrants the rescission of the contract by the plaintiff, and this does not require measuring fraudulent misrepresentation and breach according to one common scale. Example II.2 is more complex. In this case the court would have to decide whether the combination of "almost" fraudulent misrepresentation and "almost" material breach is sufficient for rescission. The court would need some common scale to evaluate both misbehaviors and aggregate them.

Or take the question, raised in our discussion of both tort law and criminal law, of whether two "almost defenses" should be sufficient to establish a valid defense (insanity and failure to mitigate in a tort case (Example I.5), or mistake and self-defense in a criminal case). Here, too, the defenses have different rationales, and any aggregation would be of a different nature from the

---

167. *See* ELIZABETH ANDERSON, VALUE IN ETHICS AND ECONOMICS 1-16 (1993) (arguing for a pluralist approach to the valuation of goods, based on the idea that goods differ in kind or quality from one another and cannot always be measured by a common criterion); Martha C. Nussbaum, *Flawed Foundations: The Philosophical Critique of (a Particular Type of) Economics*, 64 U. CHI. L. REV. 1197, 1199 (1997) (arguing that "[a] commitment to the commensurability of all an agent's ends runs very deep in the Law and Economics movement" but that it fails to describe the real world); Amartya Sen, *Plural Utility*, 81 PROC. ARISTOTELIAN SOC'Y 193, 193 (1981) (arguing that welfare economics should understand utility "primarily as a vector (with several distinct components), and only secondarily as some homogeneous magnitude").

aggregation of probabilities in the other examples that involve just factual uncertainty. To aggregate, a court would need to consider the underlying rationales directly and create a new scale that reflects the relevant theoretical considerations.

The same objection could be raised with respect to cases illustrated by our public law Example IV.3 (Targeted Killing: Two "Almost Claims"). In this example, the President claims two sources of authority for a targeted killing: a statute that gives him the authority to use military force against Al Qaeda, and his constitutional power to use military force abroad to protect American interests. Here, too, assuming neither of the legal sources independently provides the authority for targeted killing in the case at hand, a question arises of whether aggregation could lead to a different outcome. Since the rationales for the two authorizing legal sources are different, aggregation needs a common scale according to which the combined weight of the two sources as applied to the case at hand would be evaluated.

The commensurability argument might be doubted, however, because the main philosophically distinctive concern about commensurability is that values are incommensurable, and treating them as commensurable may do violence to our moral intuitions. This philosophical concern has little to do with how the law should be divided into claims, and claims into elements, which reflects institutional rather than moral considerations. It is true that comparing values of a different nature sometimes raises implementation difficulties, and those difficulties should be taken into account in considering the desirability of some kinds of aggregations. But when reduced to such a concern, the commensurability argument appears to be just another way to make the argument that aggregation has costs of implementation, which we have discussed in Section V.A above.

3. Cognitive Limitations

Aggregation requires a kind of mental manipulation that might flummox judges and juries. Consider factual aggregation. Courts do not use precise standards of proof like 40%, 51%, and 95%. Instead, the standards of proof are expressed in verbal formulations: preponderance of the evidence, beyond a reasonable doubt. Given this constraint, courts would need to give juries awkward instructions: "find in favor of the plaintiff if either claim *A* or claim *B* is supported by the preponderance of the evidence, or if claim *A* and claim *B* jointly are supported by the preponderance of the evidence even if they individually are not." In cases involving a large number of claims, the jury instructions could quickly get out of hand. But even in simple cases, like the one above, one might wonder whether juries are capable of making such fine

gradations, which would, among other things, require them to implicitly calculate joint probabilities while taking into account the degree of dependence, if any, between the two events.[168]

A similar problem could also upset normative aggregation. In Example II.2 (Non-Material Breach and Misleading Conduct), the court would instruct the jury to "find for the plaintiff if the defendant's statements were fraudulent, the defendant's breach was material, or the two actions were sufficiently serious as to warrant rescission." For Example IV.4 (Free Exercise of Religion and Free Speech), the court would find for the plaintiff if the law singled out religion and imposed a burden on it, imposed an unreasonable burden on free speech, or did not have either effect but imposed an unreasonable burden on the plaintiff's joint religion-speech rights. As in the case of factual aggregation, one might worry that juries and judges would be incapable of making the sort of fine-grained judgments that aggregation typically requires.[169]

Our response to the objections in both cases is that, while these concerns are serious, they are also marginal: the law already requires legal decisionmakers to engage in this type of mental manipulation. Legal standards require decisionmakers to aggregate factual information and normative considerations. Juries already must weigh probabilities of events and take into account the extent of dependence of events—for example, whether two witnesses who give the same testimony are entitled to extra weight because they are independent sources of information, or less weight because they might have collaborated or drawn on the same source of knowledge. Thus, while the design of legal doctrine, including the uses of aggregation, should take cognitive limitations into account, those limitations cannot by themselves provide a sufficient reason for rejecting aggregation.

---

168. *See, e.g.*, Kevin M. Clermont, *Procedure's Magical Number Three: Psychological Bases for Standards of Decision*, 72 CORNELL L. REV. 1115 (1987) (describing psychological research that suggests that people's minds process external stimuli by breaking them down into a small number of discrete categories rather than points on a probability distribution); Elisabeth Stoffelmayr & Shari Seidman Diamond, *The Conflict Between Precision and Flexibility in Explaining "Beyond a Reasonable Doubt,"* 6 PSYCHOL. PUB. POL'Y & L. 769 (2000) (discussing psychological literature on jurors' ability to distinguish standards of proof). Saul Levmore makes the interesting argument that if all (or a supermajority of) the members of the jury believe that a factual allegation is more probable than not—say, by a probability of 51%—then the probability of the allegation being true is in fact much higher than 51%. *See* Levmore, *supra* note 8, at 739-45. Levmore suggests that this could be a reason not to aggregate, mainly in cases of what we have called cross-element aggregation in this Article. *Id.*

169. The research described in Clermont, *supra* note 168, bears on this question as well.

C. Implementation

While the goal of this Article is not to offer a comprehensive proposal for courts and legislatures as to where exactly to aggregate and where not to aggregate, we raise some implementation questions and offer preliminary answers.

1. Sword and Shield

Procedurally, in most of the cases the plaintiff—or the prosecution in criminal trials—would ask the court to aggregate, because in most of the cases, aggregation would result in more liability or punishment. But as we have pointed out, aggregation could also result in less liability or punishment, depending on the circumstances. That raises the concern that plaintiffs and prosecutors may ask for aggregation only when it benefits them and avoid aggregation when it would hurt them by bringing suits in separate proceedings.

To avoid this risk, we suggest that defendants should be able to raise an aggregation defense—namely, to ask the court to aggregate claims brought against them in separate proceedings. Courts already recognize procedures for consolidating claims brought in separate proceedings, and we expect that these procedures could be adapted for cases where aggregation problems arise. In any event, we suspect that there are natural limits on such strategic behavior. For example, plaintiffs and prosecutors will often prefer to use one court rather than many courts just because it is cheaper.

2. One Trial

It is hard to imagine aggregation taking place when one claim is heard by one judge or jury and another claim is heard by another judge or jury. The main reason is the difficulty of implementation. We do not expect courts, when they aggregate, to define the exact probabilities of each claim (factual aggregation) or, alternatively, to define how close each claim is to the threshold (normative aggregation). Instead, courts should work with approximations. For example, a court should decide whether the plaintiff established, by the preponderance of the evidence, that at least one of the two claims he made is valid, or whether the aggregated normative weights of two "almost claims" he

made is large enough to justify the imposition of liability. That process of evaluation must be done by one court.[170]

But even if it were possible for courts to define the exact probabilities of each claim, or the distance between an "almost claim" and a threshold, there would be situations where aggregation between multiple courts would have more negative than positive consequences. Suppose a criminal court convicted a defendant of larceny and defined the probability of his guilt as 95% (like in Example III.2). With aggregation, and assuming that 95% is the minimum probability necessary for conviction, the defendant would have a "license" to commit another larceny (or any less severe crime) with no punishment, as long as he could be certain enough that the ex post probability of his guilt for the second larceny in the second trial would be found lower than 100%. This is so because with any probability lower than 100% for the second offense, the probability that the defendant committed the two larcenies is lower than 95%, and therefore the defendant should be punished for only one larceny.

In the preceding example, the aggregation of claims in different trials would create underdeterrence. Sometimes the reverse is true. Suppose a doctor is exonerated from liability in a medical malpractice case because the probability of his wrongdoing is 45%. Now the doctor knows that in the next case, when something goes wrong with another patient, the second patient could impose liability by showing at a very low probability that he was wronged by the doctor. Indeed, if factual aggregation across persons in such cases were allowed,[171] then the doctor would be overdeterred with respect to the second patient. But it is still possible that in the long run aggregation

---

170. Furthermore, in criminal cases, taking into account prior acquittals as a consideration for convicting the same defendant in a subsequent trial could violate the Double Jeopardy Clause. U.S. CONST. amend. V. In civil cases, aggregation across trials might seem to violate res judicata principles, but we believe it would not if aggregation in a subsequent trial were based on the exact findings of the court in the first trial—which include the probabilities of the parties' factual allegations or their normative weight—and thus would stand in no contradiction to them.

171. For aggregation across victims, see our discussion of Example I.8 (Mass Torts: Indeterminate Plaintiffs), *supra* notes 53-57 and accompanying text. Aggregation across trials creates many more implementation difficulties. In particular, in the example discussed in the text, aggregation would require apportioning the damages awarded in the subsequent trial between the plaintiffs in both trials according to criteria that would need to be set for that purpose. Aggregation across trials could be especially problematic—and objectionable—if it requires a plaintiff who won in the first trial to surrender some of his proceeds to a subsequent plaintiff. That could happen, for example, when the probability of each plaintiff's allegation being valid is higher than 50%, but the probability of both plaintiffs' allegations being valid is lower than 50%.

would do more good than bad in such cases, because it might improve deterrence with respect to the first patient.

3. Crafting Remedies

In many cases, aggregation should not require new remedies: if each claim, once accepted, would entail the same remedy, aggregation should trigger that remedy.[172]

But there are cases where aggregation requires crafting new remedies. Consider the case when each claim, if established, would have warranted a different remedy, either in magnitude (e.g., different damages awards) or in kind (e.g., damages and an injunction). Take Example I.3 (Injury in the Hospital: Two Events, Two Injuries). What amount of liability would be imposed on the hospital? At a minimum, the hospital would be liable for the less severe injury. Alternatively, the hospital could be liable for the average (or perhaps a weighted average) of the two injuries, for the more severe injury, or for a probabilistic recovery.[173] Each of these options has both advantages and disadvantages, depending on various considerations. From an efficiency perspective,[174] if courts could get all relevant information at low cost, probabilistic recovery seems to be the optimal solution.[175] If, however,

---

172. *See, e.g.*, Example II.1 (Either Material Breach or Fraudulent Misrepresentation), *supra* Section II.A.

173. Liability could be derived from the exact probabilities of the injuries being wrongfully caused by the hospital's employees. Thus, if in Example I.3 the harm to the plaintiff's leg is 100 and to the plaintiff's heart 500, liability should amount to 240: a 16% chance that both harms (100 + 500) were wrongfully caused, a 24% chance that only the harm to the leg (100) was wrongfully caused, and a 24% chance that only the harm to the heart (500) was wrongfully caused (16% * 600 + 24% * 100 + 24% * 500 = 240). If, however, we stick to the principle that liability should be imposed only for the harm that has been established by the preponderance of the evidence, the award of damages should be 100. Any other amount would not satisfy the preponderance-of-the-evidence standard. For this argument, and for an illuminating discussion of possible applications of the preponderance-of-the-evidence rule to cross-claim factual aggregation, see Alon Cohen, Implementing Aggregation in Law: The Median Outcome Procedure (Aug. 4, 2012) (unpublished manuscript) (on file with authors). Note, however, that one could make the argument that once liability is proven by the preponderance of the evidence through aggregation of claims, the *amount* of damages might be determined by a probabilistic recovery rule, even if the legal system precludes the application of this rule to one stand-alone claim.

174. Liability in this case—in any form—would probably encounter the resistance of corrective justice theorists. *See supra* notes 161-166 and accompanying text.

175. For an analysis of probabilistic recovery and its advantages and disadvantages compared to aggregation, see *supra* notes 28-32 and accompanying text.

information is costly, the average injury, or something between the two injuries, could be a better solution (the exact amount would depend on the *approximation* of the likelihoods of the two injuries being wrongfully caused by the doctors). If, however, there is a risk that aggregating the claims could result in overdeterrence or underdeterrence,[176] liability for the less severe or more severe injury, respectively, could better serve efficiency.[177]

In other cases, a plaintiff could be entitled, through aggregation, to either one type of remedy (say, damages) or another type of remedy (say, injunction), or some combination of both, while none of those entitlements, standing alone, can be established by the preponderance of the evidence. Courts could leave the choice of remedy to the plaintiff or to the defendant, or create a combination of the two remedies (for example, allowing the plaintiff to choose between damages and injunction, but a more restrictive injunction than the one he would have been entitled to if he had prevailed on his claim for injunctive relief without aggregation). These problems might seem significant, but in fact courts face similar problems under existing law, where it often occurs that plaintiffs win multiple claims that entitle them to different remedies, and courts must use discretion to choose among or combine the remedies in an appropriate manner.[178]

CONCLUSION

We have analyzed three types of aggregation in the law through various examples in four central legal fields. In most of our examples, actual courts would not aggregate. Furthermore, in most of the examples, a no-aggregation rule is taken by courts for granted, as if no other choice exists. This is puzzling, since courts sometimes do aggregate, and on many occasions the aggregation rule better serves the substantive goals of the law. In the next paragraphs we summarize our conclusions and provide recommendations for reform.

Factual aggregation. Within-element factual aggregation is routine and raises no problems. Cross-element factual aggregation is somewhat more complex, but is sometimes recognized in the law. Cross-claim factual

---

176. *See supra* note 25.

177. Note that if the probabilities in Example I.3 were 60% instead of 40%, then the plaintiff, although still allowed to recover for only one of the injuries under aggregation, would be entitled to recover for the more severe injury. Any lower recovery would lead to the absurd result that the plaintiff would be better off bringing one claim for a severe injury than bringing that same claim alongside an additional claim for a less severe injury.

178. Thus, courts routinely award both specific performance and damages in contract cases, and both an injunction and damages in nuisance cases.

aggregation can cause more serious difficulties because various and sometimes unpredictable legal consequences flow from the use of different claims. For this reason, cross-claim factual aggregation should be avoided when aggregation provides no, or minimal, benefits. We identified a few such cases: for example, when the expected liability-enhancing and liability-reducing effects of aggregation cancel out ex post, and so have no effect on ex ante incentives (assuming that the law is concerned only with those ex ante incentives).[179] Cross-claim factual aggregation should also be avoided when it involves a high level of difficulty and may confuse the jury—for example, when the probabilities are to a high degree dependent.[180]

Otherwise, courts should engage in cross-claim factual aggregation. It is important to understand that cross-claim factual aggregation does not require judges or juries to calculate probabilities with perfect precision; instead, the court should ask itself whether the probability that at least one of two (or more) claims is valid was proven at the level of proof required by the law (preponderance of the evidence, beyond a reasonable doubt, and so forth). Cross-claim factual aggregation does not require courts to do something different in nature from what they normally do—within-element aggregation. Instead of asking only whether claim *A* is more probable than not, they should also ask whether it is more probable than not that at least one of the two claims *A* and *B* holds.

Note also that cross-claim factual aggregation should not have any effect on legal norms, that is, substantive law. Its sole effect will be on the accuracy of adjudication. If courts engaged in cross-claim factual aggregation, then people would be more likely to conform their behavior to the requirements of the law.

Normative aggregation. Normative aggregation, like factual aggregation, is not always called for. We saw a number of cases where normative aggregation makes little sense because of the nature of the substantive law in question. For example, from an economic perspective, "almost negligent" behavior is actually socially desirable, so two or more instances of "almost negligent" behavior are even more socially desirable than one.[181] From a more conventional moral perspective, we suspect that two or more negligent homicides would not be considered as morally blameworthy as a single intentional homicide—although threshold deontologists might permit aggregation above a certain level.[182]

---

179. *See supra* p. 15.

180. *See supra* notes 37-39 and accompanying text.

181. *See supra* p. 22.

182. *See* EYAL ZAMIR & BARAK MEDINA, LAW, ECONOMICS, AND MORALITY 46 (2010) ("[M]oderate deontology holds that constraints have thresholds. A constraint may be

It may also be difficult for decisionmakers to distinguish fine gradations of social harm or moral wrongfulness, or even if they can, for a jury to come to a consensus about these matters. We suspect that considerations of this sort lie behind the hostility to hybrid rights in constitutional law. But, as we have emphasized, people make these sorts of judgments all the time, and expansive standard-like norms in the law require such judgments without creating insurmountable difficulties in practice.

Normative aggregation falls somewhere in the middle of the continuum between rules and standards, but it is also special because it uses existing law as building blocks rather than constructing an entirely new norm. While standards require the court to apply policy or normative considerations directly to the dispute, normative aggregation requires the court to apply existing rules, albeit in combination. The advantage of normative aggregation is that it provides more certainty and hence guidance than pure standards do.

As we have noted, normative aggregation takes two forms. First, courts might directly or explicitly aggregate claims, exemplified by the hybrid rights jurisprudence. This form of decisionmaking has an ad hoc, almost remedial quality. The plaintiff shows that she has two claims, each with normative weight that by itself is not enough to justify a remedy, but the cumulative normative weight of the two claims is enough to justify a remedy.

Second, courts might indirectly aggregate claims by using the old claims as building blocks to construct new claims. The unconscionability doctrine illustrates this process. A plaintiff has two claims under an old legal regime that are individually weak but jointly powerful; rather than give the plaintiff a remedy based on the judgment that the claims are jointly strong, the court recognizes a new claim that reflects the considerations that separately underlie the old claims. This may well be one way that the law develops over time.[183]

Cross-person aggregation. Cross-person aggregation can be both factual and normative. In many cases, the distinction is not always clear. For example, joint criminal enterprise liability enables courts to convict a defendant who, beyond a reasonable doubt, participated in crimes against at least one of a group of victims, even when the identity of her specific victims cannot be ascertained.

---

overridden for the sake of furthering good outcomes or avoiding bad ones if enough good or bad is at stake.").

183. It has been frequently argued that the law has been evolving from rules toward standards. *See, e.g.*, Peter H. Schuck, *Legal Complexity: Some Causes, Consequences, and Cures*, 42 DUKE L.J. 1, 10 (1992). If so, one part of that process may result from the type of indirect aggregation we describe, where courts replace narrow, rule-like claims with broader, standard-like claims that aggregate the policy concerns that had previously been distributed among the narrower claims.

The same doctrine also enables courts to convict a defendant of a serious crime when she, beyond a reasonable doubt, participated in minor crimes that harmed a large number of victims.

Cross-person aggregation is in tension with traditional legal doctrines that protect the rights of defendants. These doctrines are particularly strong in criminal procedure, but they exist in civil procedure as well. These doctrines are very old, and they are celebrated for, among other things, eliminating the influence of morally reprehensible norms that at one time played a significant role in legal systems — for example, holding children liable for the crimes of their parents, or members of religious or ethnic groups liable for the crimes of other members. The doctrines forced the government to show the moral culpability of the defendant, and the causal connection between her actions and the harm to a victim. But this requirement of showing the identity of the victim turns out to be a significant hurdle to justice in a world in which harms are often dispersed and their sources difficult to trace. Cross-person aggregation permits courts to overcome this hurdle without at the same time increasing the risk that people will be punished for harms that they did not cause.

We opened the Article with an observation that while individuals aggregate reasons for decisions in their daily lives, they tend *not* to expose this way of reasoning to other people. We would not be surprised if judges and jurors also aggregate more than they say they do. Courts should not only aggregate reasons when making decisions. They should make this process a more open and consistent element of adjudication.