

AGGREGATION AND LAW

Ariel Porat & Eric A. Posner*

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 each of them is too weak to justify a remedy, the plaintiff will usually lose, even if the claims are jointly powerful enough 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). (We also identify other forms of aggregation.) 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.

* Ariel Porat is Alain Poher Professor of Law at Tel Aviv University, and Fischel-Neil Distinguished Visiting Professor of Law at the University of Chicago Law School; Eric A. Posner is Kirkland and Ellis Professor of Law at the University of Chicago Law School. For very helpful comments we thank Bobby Ahdieh, Hanoch Dagan, Shai Dothan, Lee Fennell, Talia Fisher, Daniel Hemel, Aziz Huq, Roy Kreitner, Shai Lavi, Daryl Levinson, Richard McAdams, Martha Nussbaum, Adam Samaha, Adrian Vermeule, Laura Weinrib, Eyal Zamir, and participants in a faculty workshop at Emory Law School, and in the law and economics workshop at Tel Aviv University. Thanks to Lior Kotler, Ellie Norton, and Omer Yehezkel for very able research assistance.

TABLE OF CONTENTS

INTRODUCTION	2
I. TORT LAW	7
<i>A. Factual Aggregation</i>	7
<i>B. Normative Aggregation and Mixed Aggregation</i>	14
<i>C. Cross-Person Aggregation</i>	18
II. CONTRACT LAW	20
<i>A. Factual Aggregation</i>	20
<i>B. Normative Aggregation</i>	22
<i>C. Cross-Person Aggregation</i>	24
III. CRIMINAL LAW	25
<i>A. Factual Aggregation</i>	25
<i>B. Normative Aggregation</i>	30
<i>C. Cross-Person Aggregation</i>	32
IV. PUBLIC LAW	34
<i>A. Factual Aggregation</i>	34
<i>B. Normative Aggregation</i>	35
<i>C. Cross-Person Aggregation</i>	40
V. EXPLANATIONS AND PROPOSALS	42
<i>A. The Arbitrariness of Legal Boundaries</i>	42
<i>B. Other Explanations and Possible Objections to Aggregation</i>	44
1. Corrective Justice	44
2. Incommensurability	46
3. Cognitive Limitations	47
CONCLUSION	49

INTRODUCTION

Suppose you are invited to a 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 feel you 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 since 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 the 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 so they are 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. Plaintiff argues in the alternative that defendant committed a strict liability tort by driving an inherently dangerous vehicle, and caused a tort through negligent driving. To win on the strict liability claim, plaintiff must prove that the vehicle was inherently dangerous, but plaintiff can provide evidence to show only a 40 percent probability of inherent dangerousness. In addition, plaintiff can show only a 40 percent probability of negligence. A court would hold against the plaintiff because she cannot meet the 50 percent threshold for either claim. However, the plaintiff can show a 64 percent probability that the defendant committed either one tort or the other.¹ Yet a court does not permit this type of *cross-claim factual aggregation*.²

For another example, consider a plaintiff who can prove with 40 percent probability that defendant engaged in a material breach of a contract, and also can prove with 40 percent probability that defendant engaged in fraudulent misrepresentation inducing the creation of the contract. Under either theory, plaintiff would be entitled to rescission of the contract. Yet again, although the probability that at least one claim is valid is 64 percent, plaintiff would lose, because courts do not permit cross-claim factual aggregation. By contrast, *within-claim factual aggregation*, where courts simply aggregate the probabilities of the various allegations that make up a single claim, is routine.

A third type of aggregation does not require uncertainty. Suppose that plaintiff can show with certainty that defendant engaged in a minor form of fraudulent misrepresentation inducing the creation of the contract and that defendant engaged in a breach that falls just short of material. A court would typically not grant rescission. The plaintiff would lose the first claim

¹ The probability of at least one claim being valid is $1 - 0.6^2 = 0.64$.

² A variant, which we discuss *infra* text accompanying notes 22–23, is *cross-element aggregation*—where a court aggregates across elements within one cause of action.

because the fraud would be deemed mere puffery, and the second claim because the breach is not material. Yet one could argue that even if the two bad acts by defendant do not independently justify rescission, they jointly justify rescission.

Courts usually do not permit what we will call *cross-claim normative aggregation*. Yet in an important class of cases they do. 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 as well, such as the right to free speech, and yet not to a sufficient degree as to violate that right by itself, the law may nonetheless be overturned because it burdens two constitutional rights.³

One can also imagine cases that share elements of cross-claim factual aggregation and cross-claim normative aggregation. Suppose that the plaintiff can prove material breach with probability 40 percent, while the level of deception underlying the fraudulent misrepresentation claim remains a touch below what is necessary to allow rescission of the contract. 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*.

A fifth 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 percent that the pollution rather than their preexisting conditions caused 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 percent, then the residents can collectively prove that the probability that at least one of them was injured was greater than 50 percent, and therefore that the firm should pay damages (although not necessarily everyone's damages—an issue we will address later).⁴ We call this type of aggregation *cross-person aggregation*. Cross-person aggregation could be factual as in the preceding example, but also normative or mixed.⁵

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, which operate at

³ See *infra* notes 93–98 and accompanying text.

⁴ The probability that at least one resident was injured is $1 - 0.9^{10} \approx 0.65$.

⁵ The phrase “aggregation” could have various meanings in different contexts. In particular, the law often allows aggregation for evidentiary purposes: a noteworthy example is the prior acts and similar crimes doctrines applied in criminal law (*see infra* text accompanying notes 67–77), according to which past behavior of the accused could serve as evidence to prove his guilt in the present case. What is typical to this type of aggregation is the *dependence* between the accused's different misbehaviors. Our focus instead is on aggregation of *independent* claims, although we admit that sometimes the distinction between the two types of aggregation is blurred.

different levels of generality, include *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 certain types of information that is relevant to an overall evaluation of the asserted wrongdoing.

This happens in the ways we have illustrated. First, some of the factual information that is relevant for evaluating the wrongdoing of the act must be disregarded when one claim is evaluated, and other factual information must be disregarded when another claim is evaluated, even if the two claims stem from the same event. 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, 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.

Second, the law relies heavily on thresholds even when wrongdoing is typically a continuous variable.⁶ One must reach one normative threshold to show fraudulent misrepresentation and another normative threshold to show material breach. But where an event that 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.

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 harm is low-probability or, alternatively, does not quite reach the normative threshold for each individual, but across many persons, it becomes significant.

Each of these cases bears a family resemblance to each other; they all stem from the problem of aggregating two types of things: factual

⁶ Cf. John E. Coons, *Approaches to Court Imposed Compromise—The Uses of Doubt and Reason*, 58 NW. U. L. REV. 750 (1964), which argues that the law has a clear preference for “all or nothing” solutions even when those solutions are harsh, and 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.

information and normative weight. In the bulk of this Article, we will examine additional examples from torts, contracts, criminal law, and constitutional law, and then we will provide explanations and some tentative proposals for reform. Our focus will be general explanations and proposals which apply to aggregations 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 decision-makers, 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, greatly simplifies 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 and market share liability doctrines 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 the law falls short in many significant respects, some of them illustrated by our examples above. Our minimal goal is to propose “reaggregation doctrines” that permit courts to aggregate factual and normative claims where doing so does not create excessive confusion. 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 to IV analyze the non-aggregation problem in tort law, contract law, criminal law and constitutional 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.⁷

⁷ 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 22–24). Some parts of Levmore’s discussion, in particular his pointing out of implementation difficulties in aggregation, could be relevant to some types of factual aggregation that we discuss, but not to others. Another exception is Alon Harel & Ariel Porat, *Aggregating Probabilities Across Cases: Criminal Responsibility for Unspecified Offenses*, 94 MINN. L. REV. 261 (2009). Harel and Porat analyzed factual aggregation in criminal law, and focused on situations where the accused is charged with two or more separate offenses, none of them can be proven beyond reasonable doubt, but there is no reasonable doubt that he committed at least one of them. Schauer and Zeckhauser

I. TORT LAW

A. *Factual Aggregation*

Factual aggregation in tort law is common for determining whether the defendant committed a specific wrong at a given time and place. This is what we call *within-claim factual aggregation*. Thus, “if a car parked at the curb by the defendant begins to roll downhill” and hits the plaintiff, and the reason for this 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,” then the court could find the defendant liable even without knowing exactly why she was at fault.⁸

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

Example I.1. The Inherently Dangerous Vehicle. Defendant hit plaintiff while driving his vehicle. Plaintiff argues in the alternative that the vehicle was inherently dangerous, and that 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%.

Non-aggregation results in both claims being rejected. If instead the court deciding the case aggregates the two claims, it will impose liability on the defendant, since the probability that the defendant *wrongfully* hit the plaintiff is 64%, and this is more than enough for establishing liability.⁹ Courts, however, do not aggregate in cases illustrated by Example I.1.¹⁰ As

proposed aggregating probabilities across cases outside the judicial context. Frederick Schauer & Richard Zeckhauser, *On the Degree of Confidence for Adverse Decisions*, 25 J. LEGAL STUD. 27, 41–51 (1996). Schauer and Zeckhauser argued 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 separately, would not constitute sufficient reason for dismissal. Schauer and Zeckhauser maintain, however, that such an argument is inapplicable to criminal proceedings. *Id.* 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 (2002). Finally, a related literature exists on the problem of multimember decisionmaking bodies; see, e.g., 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); Jonathan Remy Nash, *A Context-Sensitive Voting Protocol Paradigm for Multimember Courts*, 56 STAN. L. REV. 75 (2003) (discussing decisionmaking by multimember judicial panels). Unlike this literature, we assume a single decisionmaker.

⁸ For this example and others, see DAN B. DOBBS, *THE LAW OF TORTS* § 154 (2000).

⁹ The probability that either claim is valid is $1 - 0.6^2 = 0.64$.

¹⁰ In *Keitz v. Commonwealth*, No. 3:11-cv-00061, LEXIS 114850 (W.D. Va. Oct. 5, 2011),

a result, defendants escape liability even when the probability that they wrongfully harmed the plaintiff is greater than 50%, just 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. Plaintiff was admitted into hospital while undergoing a heart attack. At the first stage the doctor at the emergency room misdiagnosed him and sent him home, and at the second stage, two days later, another doctor in the cardiology department gave him allegedly poor treatment during a return visit. 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 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 in different times and places and for their injurious effects the same defendant (the hospital) could be (vicariously) liable. If the two claims relating to the two events are estimated separately, liability should not be imposed; if instead the two claims are aggregated, the court should hold the defendant liable, 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.2,¹¹ but there is some lack of clarity about that. In some cases,

the plaintiff arrived at the Emergency Room (ER) of the defendant's hospital following a severe panic attack. Later he sued the hospital in tort for its failure to properly treat him at the admittance stage, based on two claims: the first that the defendant violated the Emergency Medical Treatment and Active Labor Act (EMTALA), and the second that the defendant was negligent in treating him at the ER during admittance. The court decided that both the violation of the EMTALA and the defendant's negligence were not sufficiently proven and dismissed the suit. (The negligence suit was also dismissed, for procedural reasons, but the court related to the feasibility of proving negligence despite the procedural issue). If the court had allowed aggregation of the two claims, it may have reached a different result. In *Candler General Hosp. v. McNorrill*, 354 S.E.2d 872 (Ga. 1987), the plaintiff fell and hurt his knee while being transferred by a nurse from a stretcher to a wheelchair. The plaintiff argued that the hospital was vicariously liable for the nurse's malpractice, as well as directly liable for negligence due to the inadequacy of the equipment and personnel in the Emergency Room. The court ordered that both claims should be presented before a jury, but didn't discuss aggregation of the claims. It is quite possible that the jury would dismiss the suit if both claims were examined separately, but would accept them if aggregated.

¹¹ In *Greig v. Botros*, No. 08-1181-EFM, LEXIS 25862 (D. Kan. Mar. 14, 2011), a doctor admitted a patient with chest pains to the Intensive Care Unit (ICU) of the defendant medical center without performing any tests. Six hours later, a second doctor visited the patient at the ICU, and although the patient complained that the pain was worse, the second doctor too failed to order any test. Each of the two doctors could have ordered a CT scan, which, if performed, would have saved the patient's life. With aggregation, the medical

courts were willing to impose liability on hospitals, when it was established that the plaintiff 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.¹² Furthermore, if in Example I.2 both doctors' negligence had been established by the preponderance of the evidence, and only causation had not been established by the preponderance of the evidence, it seems that courts would allow aggregation, and even impose joint and several liability on the two doctors. That latter result could be achieved if courts apply the "alternative liability rule," first established in *Summers v. Tice*,¹³ and later adopted by the Restatement (Second) of Torts. Under the alternative liability rule, as prescribed in the Restatement, "[w]here the conduct of two or more actors is tortious, and it is proved that harm has been caused to the plaintiff by only one of them, but there is uncertainty as to which one has caused it, the burden is upon each such actor to prove that he has not caused the harm."¹⁴ Indeed, the alternative liability rule has been applied to cases of one, rather than two events,¹⁵ but its logic seems to apply also to two- (or more) event

center would be vicariously liable for the wrongful death of the patient if it could be proven by the preponderance of the evidence that at least one of the two doctors was negligent in failing to order the CT scan, even if none of the claims standing alone could be proven. In *Brown v. StarMed Staffing*, 490 S.E.2d 503 (Ga. App. 1997), a patient was admitted to the emergency room at the defendant hospital, treated by a doctor and a nurse with a specific medication which was proven later to be wrong, and later discharged from the hospital by another doctor. The same day, the patient died. The court ruled that the trial court did not err in denying the hospital's motion for summary judgment. As it appears from the facts as presented by the court, there was a factual dispute regarding the negligence of each of the two doctors. Here, too, aggregation could bring about a different outcome than if each allegation with respect to each doctor were examined separately. Furthermore, the defendant hospital raised several defenses with respect to its vicarious liability for the nurse's negligence, for which aggregation could also yield a different outcome than if each defense were examined separately.

¹² In *Fieux v. Cardiovascular & Thoracic Clinic, P.C.*, 978 P.2d 429 (Or. 1999), a clamp was left behind on the plaintiff's heart during a surgery. The plaintiff could not prove who of the medical staff, composed of three nurses and one surgeon, was negligent. He relied on *res ipsa loquitur* to infer negligence. The court imposed liability on the nurses, the surgeon, and also on the hospital for vicarious liability.

¹³ *Summers v. Tice*, 199 P.2d 1 (1948) (imposing liability on two hunters for the injury one of them caused the plaintiff; while both hunters negligently shot in the plaintiff's direction, the identity of the one who actually injured the plaintiff was not established).

¹⁴ RESTATEMENT (SECOND) OF TORTS, §432B(3).

¹⁵ The wording of the Restatement, *id.*, also seems to apply for one-event cases ("Where the conduct of two or more actors is tortious, and it is proved that harm has been caused to the plaintiff *by only one of them...*") (emphasis added). It was argued by David W. Robertson, *The Common Sense of Cause in Fact*, 75 TEX. L. REV. 1765, 1787 (1997), that the same reasoning the court used in *Summers v. Tice*, *supra* note 13, was implicitly used in *Saunders Sys. Birmingham Co. v. Adams*, 117 So. 72 (Ala. 1928). In *Saunders*, two separate acts allegedly caused a road accident: one act by the car rental company, which allegedly rented the car with bad brakes, and the other act by the driver, who allegedly failed to use the brakes, or used them too late. *But see* *Smith v. Cutter Biological, Inc.*, 823 P.2d 717 (Haw. 1991) (rejecting the argument that the alternative liability rule applies when defendants did not act simultaneously); *Skipworth v. Lead Industries Ass'n*, 690 A.2d 169 (Penn. 1997) (same).

cases, as long as the negligence of each defendant can be proven by the preponderance of the evidence.¹⁶

In the next example, although not different in substance from 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 each doctor allegedly caused the plaintiff, negligently, separate harm: the doctor in the emergency room allegedly caused him a necrosis in his leg, and the doctor in the cardiology department allegedly injured his heart. The likelihood of each allegation is 40%.

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. Indeed, analogizing from the cases imposing vicarious liability on hospitals in cases similar to the one illustrated by Example I.2, and from the alternative liability rule, one could make the argument that in Example I.3 the hospital should be held liable toward the plaintiff, and if both doctors' negligence has been proven by the preponderance of the evidence, they should also be liable even if causation with respect to each of them cannot be proven by the preponderance of the evidence.

If aggregation in Example I.3 is done, 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 average¹⁷ of the two injuries, or for the more severe injury. Each of these options has both advantages and disadvantages, depending in part on the

¹⁶ See RESTATEMENT (SECOND) OF TORTS, §432B(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.” See RESTATEMENT (THIRD) OF TORTS: APPORTIONMENT OF LIABILITY, § 4 cmt. e (stating that the alternative liability doctrine is not available when the plaintiff might have caused the accident).

¹⁷ More plausibly, 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. ($16\% * 600 + 24\% * 100 + 24\% * 500 = 240$).

theory of tort law one adopts. At this stage it suffices to say that non-aggregation 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 what was *exactly* the wrongful injurious behavior which caused his harm (Example I.2), or what part of his harm is the result of that unidentified injurious behavior (Example I.3).

So far aggregations would lead to more, rather than less liability. But this is not always so. Thus, in Example I.3, with different numbers, aggregation could lead to less, rather than more liability. To see why, assume that the 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% ($60\% \times 60\%$), and 36% is not enough to establish liability.

This brings us to the interesting conclusion that when aggregation could lead to more, but also to less liability, and the injurer cannot know in advance whether in his case aggregation would lead to either the one or the other, aggregation would not necessarily change the injurer's expected liability and would 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 of each injury being caused by a doctor's negligence would be either 40% or 60% (with equal probabilities). With *no* aggregation the hospital's expected liability if the two allegations are made is 100: 50% that the probability is 40% and then it pays zero, and 50% that the probability is 60% and then it pays 200. But also *with* aggregation the hospital's expected liability is 100: 50% that the probability is 40% and then it pays 100, and 50% that the probability is 60% and then it also pays 100 (since when there are two injuries and the probability is 60%, the probability that the two injuries were caused by a doctor's negligence is only 36%). Once expected liability with or without aggregation is the same, the parties' incentives are the same as well.

Aggregation, however, would be of utmost importance for efficient incentives if the injurer could know in advance that the typical probabilities in his case would be lower 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. Conversely, if the injurer could know in advance that the typical probabilities in his case would be higher than 50%, a rule of no aggregation could¹⁸ result in overdeterrence, because under the rule of no

¹⁸ Under a negligence rule, over deterrence would not result when the standard of care was set accurately, the injurer could observe it, and the court accurately enforced it, even if the

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.

Lastly, there are cases where aggregation leads only to more liability, and then no aggregation typically leads to underdeterrence. Thus, in Examples I.1 (The Inherently Dangerous Vehicle) and I.2 (Injury in the Hospital: Two Events, One Injury), there is only one injury, when each of the two claims made by the plaintiff if properly established, justifies liability. No aggregation would allow injurers to escape liability, even if the probability that they wrongfully inflicted harm on the defendant is more probable than not.

In all the examples discussed so far, aggregation of claims would not be necessary if courts allowed probabilistic recoveries. Under a probabilistic recovery rule (PRR) 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, the PRR applies in very limited contexts (mostly in cases of lost chances of recovery¹⁹) and only when causation – but not wrongfulness – is uncertain.²⁰ It is beyond the scope of this article to discuss the advantages and disadvantages of the PRR. We draw a preliminary and brief comparison between PRR and aggregation in the footnote below.²¹

injurer pays damages higher than the harm caused by his negligence. *See* ROBERT COOTER & THOMAS ULEN, *LAW AND ECONOMICS* 218–9 (2012) (arguing that a small change in the damages awarded to the victim will not cause the injurer's behavior to change).

¹⁹ *See e.g.*, *Herskovits v. Group Health Coop. of Puget Sound*, 664 P.2d 474, 476–77 (Wash. 1983) (holding that a 14% reduction, from 39% to 25%, in the decedent's chance for survival was sufficient evidence to allow the case to go to the jury). Some courts have adopted the lost chance doctrine only in cases of the victim's demise, rejecting it in other cases. *See Falcon*, 462 N.W.2d at 57–58 (allowing probabilistic recovery only where the ultimate harm to the victim is death). For straightforward support of applying a probabilistic rule to lost chance of recovery cases, see Joseph H. King, Jr., *Causation, Valuation, and Chance in Personal Injury Torts Involving Preexisting Conditions and Future Consequences*, 90 *YALE L.J.* 1353 (1981). *See also Doll v. Brown*, 75 F.3d 1200, 1206–07 (7th Cir. 1996). In *Doll*, Judge Posner supported extending the lost chance principle to areas beyond medical malpractice. Specifically, he instructed the lower court to consider the possibility of awarding the plaintiff in an employment discrimination suit damages calculated according to the chances that his not being promoted was due to illegal discrimination.

²⁰ *See* ARIEL PORAT & ALEX STEIN, *TORT LIABILITY UNDER UNCERTAINTY* 57–83 (2001) (discussing the pros and cons of probabilistic recoveries in different contexts, including cases where wrongfulness is uncertain).

²¹ 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 to all claims. Second, to apply PRR courts need accurate information about 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. With aggregation, courts do not need to calculate exact probabilities: they only need to determine whether after aggregation it is more probable than not that the defendant wronged the plaintiff (or that a defense applies), as they do also when there is only one

The last example in this section, which was first analyzed by Saul Levmore,²² addresses a case where factual aggregation occurs within a cause of action but across elements—which we will call *cross-element aggregation*, and it typically²³ leads to less rather than more liability.

Example 1.4. Several Elements of One Cause of Action. Plaintiff argues that defendant was negligent and that that negligence is the cause of his injury. The probability that defendant was negligent is 60% and the probability that defendant, given his negligence, caused the harm, is also 60%.

With no aggregation the court would find the defendant liable, and with aggregation liability would not be imposed. Specifically, aggregation would yield that the probability that the defendant negligently caused the litigated harm is only 36%, and under the preponderance of the evidence rule that probability is too low for imposing liability. Note that the aggregation problem becomes more acute as the number of elements composing the claim raises. Thus, if in addition to the uncertainty with respect to negligence and causation there is also uncertainty with respect to the plaintiff's harm, so that each of the three elements is proven at probability of 60%, aggregation would yield probability of 21.6% that the defendant negligently caused the litigated harm. The law is not clear as to whether jury and judges should engage in cross-element aggregation: in several jurisdictions jury instructions encourage such aggregation, while in other jurisdictions they discourage it.²⁴

claim, and cross-claim factual aggregation is not done. Third, corrective justice theorists resist PRR because PRR makes 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 also raise the 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. Aggregation would therefore be easier to accept for many who resist PRR. See discussion *infra* text accompanying note 123. There are also advantages to PRR 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 aggregation rule would not take place under PRR.

²² Levmore, *supra* note 7, at 723, 725–8.

²³ It would lead to more liability if some elements were alternatives to one another. Levmore, *id.* at 726–9.

²⁴ Levmore, *id.* at 752 n. 58 (arguing that no jurisdiction explicitly recognizes cross-element aggregation), and at 725, 741 (arguing that jury instructions tend to be ambiguous in several states, implicitly allowing such 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); Ronald J. Allen & Sarah A. Jehl, *Burdens of Persuasion in Civil Cases: Algorithms v. Explanations*, 2003 MICH. ST. L. REV. 893 (2003) (arguing that the jury is never instructed to do what we call cross-element aggregation).

Throughout this section we have assumed that the probability of the two (or more) claims being valid is 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.²⁵ 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 Injuries), the validity of the claim that the defendant negligently caused the first injury could increase the probability of the validity of the claim that the defendant negligently caused the second injury, and vice versa. This complication, however, does not preclude aggregation.

To illustrate, assume that claim A's probability standing alone is 40%, and claim B's probability standing alone is also 40%, as in Example I.3. Assume now that the probabilities of the two claims are dependent, and because of that, the probability of each claim increases to 50%. With no aggregation, both claims will be dismissed since a 50% probability is not enough to establish liability. With aggregation, however, one claim would be accepted, since the probability that at least one claim holds is greater than 50%. Indeed, because the probabilities of the two claims are dependent, the probability that at least one claim is valid is less than 75%, which would have been the result of aggregation if both claims (with 50% 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 wrong, claim B also is wrong 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.

B. Normative Aggregation and Mixed Aggregation

Consider the following example.

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

²⁵ Levmore, *supra* note 7, at 726–8 (discussing the dependence problem mainly in cases represented by our Example I.4).

applies.²⁶ 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.²⁷

The court deciding the case would not aggregate the two defenses raised by the defendant, and would reject both of them.²⁸ 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 to justify a defense the defendant should reach a point of normative weight denoted as *a*. That point can be reached if one of the two defenses applies (thus, each defense 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). Consider this argument from an economic perspective. We 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 want to deny damages where the barely mentally competent person will be only barely responsive to them *and* the surgery-fearing victim will be 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 if damages are denied.

A 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. By doing so the court would acknowledge that even if none of the defenses *standing alone* should apply, the accumulative weight of the two "almost defenses" is sufficient for

²⁶ The insanity defense is quite limited under American tort law: *See* DOBBS, *supra* note 8, at §120 (stating that the general rule is that the mentally disabled are liable for negligence, but a few exceptions exist: if the act was committed as a result of a sudden onset of unforeseeable insanity, or the insanity was caused by the defendant's effort to protect the plaintiff, or the plaintiff's job involved working with mentally disabled, some jurisdictions will not impose liability); *Breuning v. American 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, liability will not be imposed).

²⁷ *See* RICHARD EPSTEIN, TORTS 448 (1999) (discussing the dilemma of whether refusing to undergo a surgery with a positive expected benefit would necessarily infringe the mitigation of damages principle).

²⁸ In *Boa v. San Francisco-Oakland T. Rys*, 187 P. 2 (Cal. 1920), the plaintiff sustained physical injury when she exited from the defendant's street car. The defendant argued that the plaintiff was contributorily negligent, and that she failed to mitigate damages by choosing an improper physician. The trial court instructed the jury that if they found the plaintiff's contributory negligence to be 50% or less, they should not allow the contributory negligence defense. 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.

justifying a defense.

As we have said, courts would probably not allow aggregation in example I.5, but maybe would be more attentive to aggregation arguments when facing two defense arguments based on similar normative grounds, or when the defenses interrelate with each other. Thus, if a defendant raises the defenses of self-defense and insanity, arguing that the somewhat excessive force she used to defend herself is related to her deficient mental capacity at the time of the injury, maybe the justification to aggregate the two “almost defenses” would make more sense to some courts.²⁹

However, courts should be cautious with cross-claim normative aggregation, because the weight of “almost” defense, or “almost” claim could be zero, and then there would be nothing to aggregate. For example, suppose a driver hits a pedestrian and then subsequently crashes into the pedestrian’s house. A court holds that each act was almost negligent 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 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 for negligence. If, for example, we believe that there is some moral blame in causing harm even if non-negligently, but that the level of blame by itself is not enough to justify the law’s intervention, then we might believe that that latter conclusion should change 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 ultra-hazardous activities: even if an injurer’s activity is efficient, the high intensity of creating risks to victims is the justification for imposing liability on him.³⁰

An example where we believe that cross-claim normative aggregation could explain a puzzling legal rule in tort law, is the rule that one has a duty to rescue another person under the common law, if the first person was the one who created the risk (even not negligently) to the person needing the rescue. Thus, a person who non-negligently shot her gun in the forest and hit the plaintiff causing him to fall into a pool of water, must take

²⁹ In a criminal context, see *State v. Peterson*, 857 A.2d 1132 (Md. App. 2004), where a defendant was convicted of murder in the first degree of her husband. On appeal, the defendant argued that the trial court erred when it rejected both her self-defense and insanity arguments. She argued that since she was a battered wife, she used excessive force defending herself, but the excessiveness of the force itself was the result of insanity, since she suffered from “battered spouse syndrome.” The court accepted the validity of her arguments and ordered a new trial.

³⁰ Cf. Gregory C. Keating, *Reasonableness and Rationality in Negligence Theory*, 48 STAN. L. REV. 311 (1996) (arguing that when risks are high, prevention is justified, even if cost-benefit analysis yields that it is not); Gregory Keating, *Distributive and Corrective Justice in the Tort Law of Accidents*, 74 S. CAL. L. REV. 193 (2000) (same).

reasonable measures to rescue him, although other people do not have such a duty.³¹ Tort theorists struggled with the question of what the justification for such a rule is, given that both non-negligently causing harm and non-rescue do not give rise to tort liability. Isn't it that nil plus nil is still nil?³² 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 two claims are aggregated, liability is justified. In particular, non-negligently causing harm is not sufficient to justify liability, and a failing to rescue is not sufficient to justify liability, but non-negligently causing harm followed by a failure to rescue may nevertheless justify liability.³³

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 as to the application of the insanity defense, so that the probability that that defense applies is 40%. If, however, the uncertain facts reflected reality, the defense would clearly apply.

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 will also reject the mitigation of damages defense, since the failure of the plaintiff to undergo surgery, even

³¹See EPSTEIN, *supra* note 27, 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); DOBBS, *supra* note 8, at §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); *Maladona v. Southern Pac. Trans. Co.*, 629 P.2d 1001 (Ariz. 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).

³²See Richard A. Epstein, *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).

³³This is not the only possible explanation, of course. See EPSTEIN, *supra* note 27, 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); Ernest J. Weinrib, *A Call for the Duty to Rescue*, 90 YALE L. J. 247, 257–8 (1981) (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 has a duty to act); William Lands & Richard Posner, *Salvors, Finders, Good Samaritans and Other Rescuers: An Economic Study of Law and Altruism*, 7 J. LEGAL. STUD. 83, 125–6 (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).

if considered by most people unreasonable, does not trigger the application of the defense.³⁴ Conversely, with aggregation the court would recognize a defense for the plaintiff and would exonerate her from liability from the harm she could have mitigated if she had undergone the surgery.³⁵

C. Cross-Person Aggregation

Should tort law allow cross-persons aggregations? Take the following example:

Example 1.7: Mass Torts: Indeterminate Plaintiffs. Defendant's factory wrongfully emits radiation which caused an increase in the frequency of a fatal cancer in the population; instead of 100 people contracting the disease every year, now 125 people contracting it every year. Due to lack of scientific knowledge it is impossible to identify who are those 25 victims whose disease was caused by the radiation. All 125 people bring suits against Defendant.³⁶

Under traditional causation principles all suits would be dismissed because none of the plaintiffs can establish by the preponderance of the evidence her claim 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 would impose liability on defendant for 20% of the total harm suffered by all plaintiffs and would distribute the damages among them (in equal shares, if all suffer the same harm).

³⁴ 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. 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. The plaintiff argued that he was not contributorily negligent and that when he signed the compromise agreement he was under the influence of drugs. 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. If the defense relating to the release agreement was "almost" applicable, a mixed aggregation could have brought the court to a different decision.

³⁵ As with factual aggregation, the probabilistic recovery rule 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* note 21 and accompanying text. For a recent discussion of the related topic of the either/or character of the law, see Leo Katz, *WHY THE LAW IS SO PERVERSE* 139–81 (2011), who also cites the main sources in the literature.

³⁶ For a well-known case presenting the same 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) (Suits brought by veterans against manufacturers of "Agent Orange" for injuries allegedly caused by their exposure to that chemical, which was used by US military forces in the Vietnam War. Many of the plaintiffs' injuries could have equally been attributed to either their pre-existing conditions or to the exposure to Agent Orange).

Notice that the aggregation of all claims in Example I.7 mimics defendant's liability if there was no uncertainty, but does not mimic plaintiffs' entitlements if there was no uncertainty. Without uncertainty defendant would probably have paid about 20% of the total harm as he also does under uncertainty, but damages go to 25, rather than 125 plaintiffs. Thus, one could say that the aggregation works on defendant's rather than the individual plaintiffs' side.

Market Share Liability is an example of aggregating on both defendants' and plaintiffs' sides, and therefore mimics both defendants' liability and plaintiffs' entitlements if there was no uncertainty. Thus, in the notorious DES cases, numerous manufacturers produced the same generic drug for preventing miscarriage, which, many years later, was proven to be defective and harmful to the daughters of the women who had taken the drug.³⁷ 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.³⁸ For some time, courts 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%.³⁹ In 1980, in *Sindell v. Abbott Laboratories*, the California Supreme Court established the Market Share Liability doctrine, whereby all manufacturers are liable toward plaintiffs in accordance with their market share.⁴⁰ Under market share liability, when all suits are completed, manufacturers bear liability in the amount of the actual harm they wrongfully caused and plaintiffs receive damages in the amount of the harms they suffered from wrongdoing. Thus market share liability aims at mimicking the world without uncertainty, and aggregation works on both defendants' and plaintiffs' sides.

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

In *U.S. v. Hatahley*,⁴¹ Native Americans brought suit against the US 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

³⁷ *Sindell v. Abbott Labs.*, 607 P.2d 924 (Cal. 1980) (establishing the Market Share Liability doctrine and elaborating on its advantages).

³⁸ *Id.* at 928.

³⁹ *See, e.g.,* *McCreery v. Eli Lilly & Co.*, 150 Cal. Rptr. 730, 732–35 (Cal. Ct. App. 1978) (denying liability of 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).

⁴⁰ *Sindell*, 607 P.2d at 937–38.

⁴¹ 257 F.2d 920 (10th Cir. 1958).

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.”⁴² That theory allowed the court to be generous to the plaintiffs and award them relatively high amount of damages. The Court of Appeals 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.”⁴³ Using our terminology, the district court allowed a kind 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.⁴⁴

II. CONTRACT LAW

A. *Factual Aggregation*

Like tort law, contract law permits factual aggregation within claims,⁴⁵ but does not generally permit cross-claim factual aggregation.

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

Although the probability that one claim or the other is valid is 64 percent, plaintiff would lose, because courts do not permit cross-claim factual aggregation.⁴⁶

⁴² “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 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. It was a community loss and a community sorrow shared by all.” *Id.* at 17.

⁴³ The court reversed and remanded the case for a new trial as to damages. *Id.* at 15.

⁴⁴ Alternatively, maybe the district court thought that the cross-person aggregation allowed the award of higher total damages than the total damages that would have been awarded if damages for the emotional pain and suffering had been calculated for each plaintiff in isolation.

⁴⁵ *See e.g.*, *Photovest Corporation v. Fotomat Corporation*, 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). However, it is not clear whether the individual allegations were below the preponderance of the evidence standard and were aggregated, or were each above the preponderance of the evidence standard. If the latter is right, than it is an example of a normative aggregation.

⁴⁶ *See, e.g.*, *Zytax, Inc. v. Green Plains Renewable Energy, Inc.*, No. H-09-2582, 2010 WL

Long-term contracts or business relationships involving multiple contracts can raise issues of cross-claim factual aggregation. Thus, buyer could bring a suit against seller arguing that seller breached the same contract several times in the past, or breached several contracts in the past, and therefore is entitled to compensation. Buyer may fail, however, to prove any specific breach by the preponderance of the evidence, so the court will dismiss the suit. By contrast, aggregation would lead the court to award damages for one or some of the alleged breaches, or, in the appropriate cases, to confirm that buyer's refusal to offer payments for performance was justified in one, or more, of the alleged breaches. Indeed, the court would not be able to point out the exact breach that took place, and therefore the court would have to craft a remedy that averages the alleged breaches. If the alleged breaches are similar (say, five deliveries of the same amount of widgets, with similar allegations of breach), aggregation would be relatively easy. Otherwise, aggregation would be more complex and one could argue that aggregation should not be made. But even if averaging the remedy is complex, at a minimum the court should acknowledge that the least severe breach took place and allow a remedy for it.

It is possible that courts in fact make such aggregations without admitting it. Moreover, sometimes there is dependence between the alleged breaches, so that one alleged breach, even if not proven by the preponderance of the evidence, could serve as evidence to establish other breaches.⁴⁷ Thus, similar to the criminal law doctrines of prior acts and similar crimes,⁴⁸ one alleged breach could enhance the probability that another alleged breach took place.

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, they do not.⁴⁹ This 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 does not improve incentives because the too-high and too-low outcomes cancel out

4702303 (S.D.Tex. Nov. 10, 2010), where the defendant tried to avoid a contract by alleging fraudulent misrepresentation, material breach, and mutual mistake, all in the alternative. The court dismissed each claim separately and did not treat them in the aggregate.

⁴⁷ See *Igen-International v. Roche Diagnostics GMBH*, 335 F.3d 303 (4th Cir. 2003) (affirming the jury's aggregation of several minor breaches to form a material breach).

⁴⁸ See *infra* text accompanying notes 67–77.

⁴⁹ 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, best efforts, good faith, etc. They also often create mechanisms for the resolution of disagreements by non-lawyer arbiters who need not provide rigorous and formal reasoning for their decisions. Those standards and mechanisms could be used for implicit aggregations.

ex ante.

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% (with equal probabilities). With *no* aggregation the promisor's expected liability if the two allegations are made is 100: 50% that the probability is 40% and then he pays zero, and 50% that the probability is 60% and then he pays 200. But also *with* aggregation the promisor's expected liability is 100: 50% that the probability is 40% and then he pays 100, and 50% that the probability is 60% and then he also pays 100 (since when there are two breaches and the probability is 60%, the probability that the two breaches took place is only 36%). Once expected liability with or without aggregation is the same, the parties' incentives are the same as well.

If, however, the remedy the parties to the contract anticipate is rescission, and the two alleged breaches are material (or there are an alleged material breach and an alleged fraudulent misrepresentation, as in Example II.1), aggregation would *always* increase the promisor's ability to rescind the contract. That would generally improve the parties' incentives, 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 Minor Fraudulent Misrepresentation. Defendant engages in a minor form of fraudulent misrepresentation in order to secure plaintiff's consent to a contract; subsequently, defendant engages in a breach that falls just short of material. Plaintiff seeks to rescind the contract based on both fraudulent misrepresentation and breach.

A court would typically not approve the rescission in Example II.2.⁵⁰ Plaintiff would lose the first claim because the fraud would be deemed mere puffery, and on the second claim because the breach is not material. Therefore, the court would decide that the plaintiff, by unlawfully

⁵⁰ At least not expressly. Courts sometimes permit considerations from one claim (or defense) to bleed over to another. For example, in *Lincoln Ben Life Co. v. Edwards*, 45 F.Supp.2d 722 (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.

rescinding the contract, breached the contract himself. By contrast, 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 none of them standing alone was sufficient for rescission.

Yet aggregation may be permitted when within a claim.

Example II.3. Two minor breaches. Defendant promises to build a house for plaintiff. When the time for the first progress payment comes round, defendant is a little behind in schedule and has made some minor mistakes in construction.

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 plaintiff.⁵¹

Consider another example, which shows the evolution of the law to address aggregation problems.

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 customer defaults on payments for the TV set.

Under older doctrine, the customer 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 treat lack of education as falling short of incompetence or undue influence.⁵² But over the last half century, the doctrine of unconscionability evolved.⁵³ Under this doctrine, plaintiff can

⁵¹ See *Seven-Up Bottling Co. (Bangkok) v. PepsiCo Inc.*, 686 F.Supp 1015 (S.D.N.Y. 1988) (multiple breaches by bottler; similar analysis); *SunTrust Mortg., Inc. v. United Guar. Residential Ins. Co.*, No. 3:09cv529, 2011 WL 3664749 (E.D.Va. Aug. 19, 2011) (holding that multiple small breaches added up to material breach justifying rescission); *Clarendon America Ins. Co. v. General Sec. Indem. Co. of Arizona*, 124 Cal. Rptr. 3d 1 (Cal. App. 2011) (allowing the insurer not to pay the contractor for property damages, probably because of the cumulative weight of two defaults of the insured under the policy).

⁵² *Mason v. Acceptance Loan Co.*, 850 So. 2d 289 (Ala. 2002) (stating that lack of education does not deprive one of the ability to contract, and that not reading the contract, in the absence of fraud, is not a reason to avoid the contract).

⁵³ See *Williams v. Walker-Thomas Furniture Co.*, 350 F.2d 445 (D.C. 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); UCC §2-302(1) (“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

jointly invoke considerations—mistake, lack of education—that can be considered only individually under other doctrines. In this way, the unconscionability doctrine can be understood as permitting a kind of aggregation. But it is important to see that the aggregation is indirect: the court does not say that plaintiff can prevail by presenting colorable claims under two doctrines; instead, it creates a new doctrine that has the same effect.⁵⁴

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 while reflecting 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.⁵⁵ We will return to this problem in Part V.

C. Cross-Person Aggregation

As with torts also with contracts there are cases where plaintiffs cannot prevail against a specific defendant because of inherent difficulties of proof and then a question could arise whether cross-person aggregation should be allowed. The next example illustrates such cases.

Example II.5. Many Unproven Breaches with Customers. Defendant ships goods by sea, and plaintiffs are 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 defendant, but that evidence is not strong enough to establish liability.

The court in Example II.5 would reject all suits because in none of them the plaintiff can establish liability. Prevailing law would also not allow recovery even if plaintiffs brought a class action, since in order to succeed in a class action plaintiffs must show that they would have succeeded in trial if they had brought their claims separately.

That result would change if courts were willing to aggregate all claims

the unconscionable clause, or it may so limit the application of any unconscionable clause as to avoid any unconscionable result.”); see JAMES J. WHITE & ROBERT S. SUMMERS, UNIFORM COMMERCIAL CODE 151–177 (2000) (discussing the unconscionability doctrine); E. ALLAN FARNSWORTH, FARNSWORTH ON CONTRACTS 908 (1999) (same).

⁵⁴ A similar point can be made about *Odorizzi v. Bloomfield Sch. Dist.*, 54 Cal. Rptr. 533 (Cal. App. 1996), where fraud, duress, and similar claims were rejected, but an undue influence claim which took into account factors relevant to the other claims was accepted.

⁵⁵ For this criticism of the unconscionability doctrine, see, e.g., Evelyn L. Brown, *Uncertainty of U.C.C. Section 2-302: Why Unconscionability Has Become a Relic*, 105 COM. L. J. 287, 288 (2000) (arguing that the unconscionability doctrine allowed courts wide latitude in its use as well as manipulation of the term, and increased the potential for arbitrary decisions).

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 was established by the preponderance of the evidence and only causation could not be proven, than if the breach itself cannot be proven. By analogy from tort law, when the wrongdoing is proven but causation is uncertain the law is more lenient to plaintiffs, than if the wrongdoing itself cannot be proven.⁵⁶

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.⁵⁷ Consider the following example:

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%.⁵⁸

Under prevailing law the defendant would be acquitted of both offenses. Yet, there is a 99% probability⁵⁹ 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 would convict the defendant of one unspecified offense and impose on him *at a minimum* the sanction of the less severe of the two offenses, that is, pickpocketing. 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

⁵⁶ See discussion *supra* text accompanying notes 12–14.

⁵⁷ See Harel & Porat, *supra* note 7 (discussing factual aggregation in criminal law).

⁵⁸ This example is borrowed from Harel & Porat, *id.*, at 262.

⁵⁹ The probability that the defendant committed each one of the offenses is .9, and therefore the probability, for each one, that he did not commit the offense is $1 - .9 = .1$. Consequently, the probability that he did not commit any offense is $(.1)^2 = .01$, and the probability that he committed at least one of the offenses is $1 - .01 = .99$.

convicted.

Example III.1 illustrates how cross-claim factual aggregation can result in more convictions than with no aggregation. But aggregation can 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 any 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 on both charges because the probability that he committed each of the offenses (95%) is sufficient for conviction. Yet, the probability that the defendant committed *both* offenses is only 90%, which is lower than 95%.⁶⁰ 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% (it is 99.75%) which is sufficient for conviction, the probability that he committed two offenses is lower than 95% which is insufficient for conviction. The court would need to decide which of the two offenses to convict the defendant of; the correct decision is to convict the defendant only of the most severe offense since the probability that he committed that offense is 95%.⁶¹ Put differently, the probability that the defendant deserves the combined sentence for both offenses is less than 95%, but the probability that he deserves the sentence for the more severe offense is higher than 95%. However, it could well be appropriate for the court to nudge the sentence up a bit to reflect the fact that there is a 90% probability that the defendant committed both offenses.

Cross-claim factual aggregation in cases illustrated by Example III.1 would improve deterrence (assuming, as we must, that the beyond-the-reasonable-doubt standard should be taken as fixed). Under current law, defendants who are charged with several offenses and the probability of their guilt is very high often escape conviction just because no *specified* offense can be attributed to them. Those defendants are underdeterred under current law, and will be better deterred with cross-claim factual aggregation. In contrast, cross-claim factual aggregation in cases illustrated by Example III.2 would reduce wrongful convictions. Under current law defendants who are charged with several offenses are often convicted of

⁶⁰ $(.95)^2 = .9025$.

⁶¹ Cross-claim factual aggregation could apply also to defense claims. In *Ralston v. State*, No. 49A02-0909-CR-929, LEXIS 693 (Ind. App. May 26, 2010), the defendant raised two defense claims: that he did not cause the victim's death and that he acted in self-defense when he repeatedly punched the victim. Both claims were considered separately by the jury and denied. Aggregating the two claims might have brought a different result.

those offenses, even if the probability that they committed *all* those offenses is too low. With cross-claim factual aggregation those defendants will be convicted of fewer offenses and many wrong convictions will be avoided.⁶²

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. 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 nature against the accused at the same trial, since it is hard to imagine that aggregation would take place when each charge is brought before a different jury or judge.⁶³ 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. However, from a different perspective the lower costs of bringing several low-probability charges indicate an advantage of aggregation: it economizes on enforcement costs. 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—so his criminal record would literally list his offense as “rape or larceny”—and some would consider that intolerable, although it is hard to imagine why acquittal of both crimes would be preferable.⁶⁴

Cross-element aggregation is also an issue in criminal law.⁶⁵ If several

⁶² Under certain conditions, reducing the number of wrong 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 sanction of the guilty and innocent. Cf. Chris William Sanchirico, *Character Evidence and the Object of Trial*, 101 COLUM. L. REV. 1227, 1276 (2001) (arguing that if bad character evidence were admitted at the conviction stage, the disincentive for engaging in crime would be weakened, since character evidence enhances the probability of conviction, both for those who committed the proscribed acts and for those who refrained from such behavior, leading to a decrease in the marginal cost of engaging in the criminal activity *ex ante*).

⁶³ 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. 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.* 14(a).

⁶⁴ For more objections and responses see Harel & Porat, *supra* note 7, at 291–309.

⁶⁵ See Maya Bar-Hillel, *Probabilistic Analysis in Legal Factfinding*, 56 ACTA PSYCHOLOGICA 268–70, 282–83 (1984) (analyzing the use of probabilities in cases and suggesting a “soft role... for probability in the *factfinding* process”). *But see* *People v. Collins*, 438 P.2d 33, 33, 40 (Cal. 1968) (rejecting the use of probabilities in determining

elements of the same offense must be proven to establish the defendant's guilt, then cross-element aggregation generates a different outcome than if each element is 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.⁶⁶

Cross-claim factual aggregation should be distinguished from two existing doctrines in criminal law: the prior-acts and similar-crimes doctrines.⁶⁷ Under both of these doctrines, past similar behavior on the part of the defendant can be used as evidence supporting conviction.⁶⁸ 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 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,⁶⁹ 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, or absence of mistake or accident.⁷⁰ This evidence cannot be used to prove the defendant's bad character and courts are required to instruct the jury accordingly.⁷¹ Interestingly, under Rule 404(b), as interpreted by

guilt on the facts of this case).

⁶⁶ Cf. Levmore, *supra* note 7, 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 7, at 138 (discussing the rule of aggregation in the context of voting by judges in a panel or by jurors and observing that "[a]lthough a criminal defendant cannot be convicted unless a jury unanimously finds each element of the crime charged proven beyond a reasonable doubt, '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'" (citation omitted)). Note that the Model Penal Code section 1.12(1) says that "[n]o person may be convicted of an offense unless each element of such offense is proved beyond a reasonable doubt." It seems that the code rules out cross-element aggregation.

⁶⁷ See FED. R. EVID. 403, 413, 414.

⁶⁸ *Id.*

⁶⁹ See *id.* 404(b). See also *United States v. Woods*, 484 F.2d 127, 137 (4th Cir. 1973), where the court stated prior to the enactment of Rule 404(b): "Unlike other cases where evidence of prior crimes is admissible for only limited purposes and where it is necessary or proper to give limiting instructions, evidence of the prior events was admissible here to prove both that Paul was the victim of infanticide and that defendant was the perpetrator of the crime."

⁷⁰ FED. R. EVID. 404(b).

⁷¹ See *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 MRE 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

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.⁷²

The similar-crimes doctrine, adopted in Rules 413 and 414 of the Federal Rules of Evidence, applies to sexual assault and child molestation offenses.⁷³ Under this doctrine, if the defendant is accused of one of these types of offenses, “evidence of the defendant’s commission of another offense or offenses of sexual assault or child molestation is admissible, and may be considered for its bearing on any other matter to which it is relevant.”⁷⁴

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 influences the likelihood of conviction.⁷⁵ 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 is more likely to either have intended or have 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.⁷⁶ 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). The cross-claim factual aggregation is not based on any dependence between the offenses attributed to the defendant: the probability that she committed one offense does not change the probability that she committed another. Rather, only the probability that

guilt.”); *see also* *Huddleston v. United States*, 485 U.S. 681, 689–92 (1988) (holding that the trial court is not required to make a preliminary finding that the petitioner proved commission of the similar acts by a preponderance of the evidence).

⁷² *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).

⁷³ FED. R. EVID. 413, 414.

⁷⁴ *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 (1997) (“By requiring the admission of propensity evidence, the rules prevent a fundamentally fair trial, and thus violate due process...”).

⁷⁵ As Example III.2 illustrates, sometimes aggregation leads to acquittal rather than to conviction.

⁷⁶ In probability theory, two events are dependent if the probability of one is a function of the occurrence of the other; otherwise, they are independent.

she committed *an unspecified offense* is affected.⁷⁷

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 for exonerating the defendant from liability; but the weight of the two “almost” defenses taken together may be sufficient for acquittal. For example, the defendant may raise a factual mistake defense and self-defense. Suppose that neither of the two defenses reaches the point where it applies – the mistake was unreasonable (but not extremely unreasonable) and the defendant used unreasonable (but not extremely unreasonable) force.⁷⁸ 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 interesting, cross-claim normative aggregation could be made also across cases, 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 Non-Criminal Misdemeanors. Defendant is accused of five separate offenses, allegedly committed in different times and places, of interrupting with the work of a public official. Each behavior considered separately does not reach the point where the behavior is defined as an offense.

Under prevailing law the defendant would be acquitted of the five

⁷⁷ 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. Under cross-claim factual aggregation, in contrast, as illustrated by Example III.2, the fact that a person was convicted of several offenses in the past decreases the probability of conviction in a later case.

⁷⁸ In *State v. Thornton*, 730 S.W.2d 309 (Tenn. 1987), the defendant raised self-defense and heat of passion provocation mitigation. He lost both defenses at trial, and was convicted of murder in the first degree, but won the mitigation as a matter of law on appeal, where the court reduced his conviction to voluntary manslaughter. It is doubtful whether he deserved that defense, so a possible explanation for the decision is that the court implicitly engaged in normative aggregation of two “almost” defenses. In *Wesley Johnson v. State*, 36 So. 3d 170 (Fla. App. 2010), the defendant raised two defense claims: that he was not present at the place where the murder took place, and that the co-defendant’s fatal beating of the victim was an unforeseen independent act falling outside of the original plan of the crime. The jury denied both claims and convicted him of second-degree murder. The court affirmed the trial court’s instructions to the jury. From the facts as presented by the court of appeals, it is possible that there was one defense claim that was not established factually (the causation claim) and another defense claim (unforeseen independent act) that was, if at all, an “almost” defense claim. If this reading of the case is right, then a mixed aggregation (factual and normative) could have brought a different result.

charges brought against him.⁷⁹ But if all cases are 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 could think of two main reasons why five cases could justify conviction even if one case does not: first, maybe 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 defendant's behavior as a whole and may justify convicting him.

However, criminal law has a number of aggregation doctrines which allow courts to 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, one single behavior of stalking does not constitute an offense, but if that behavior occurs several times, then at a certain point it becomes an offense.⁸⁰ The New Jersey Criminal Code, for example, defines stalking as "repeatedly maintaining a visual or physical proximity to a person...."⁸¹ Thus, normative aggregation is allowed under the Code: a possible interpretation of the Code is that one act of stalking is bad, but not bad enough to justify the law's intervention. A second interpretation is that one act of stalking could be accidental with no malicious motives, and therefore, in order to reduce the risk of false convictions, the law requires for conviction more than one act of stalking.⁸²

In other cases, a single behavior is an offense, but if that behavior is repeated several times the series of offenses could 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

⁷⁹ See e.g., *State v. Stolen*, 755 N.W.2d (Neb. 2008) (reversing the accused's conviction on the ground that physical interference is required in order to commit obstruction of government operations); *Ovide Duncantell v. State*, 230 S.W.3d 835 (Tex. App. 2007) (The defendant was convicted of interference with the duties of a public servant, since he disregarded the officer's repeated request to stand back. The court decided that by repeatedly entering the crime scene area, the defendant interfered with the officer's duty).

⁸⁰ See Heather C. Melton, *Stalking: A Review of the Literature and Direction for the Future*, 25 CRIM. JUST. REV. 246, 247 (2000) (stating that "[t]he term 'stalking' is used to describe the willful, repeated, and malicious following, harassing, or threatening of another person")

⁸¹ N.J. Stat. § 2C:12-10-a. For the New Jersey Superior Court's interpretation of the stalking clause, see *State v. Cardell*, 723 A. 2d 111 (N.J. Super. 1999).

⁸² In *State v. Berg*, 213 P.3d 1249 (Ore. 2009), the defendant was convicted with 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. 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.

possession of the first degree; if however the quantity is large enough, then he would be convicted of the offense of a higher degree. The large quantity condition could be satisfied even if the accused imports drugs several times, each time the quantity is small, but the total quantity across all occasions is large enough.⁸³

Another example of normative aggregation is the doctrine embedded in the Racketeer Influenced and Corrupt Organization Act (RICO).⁸⁴ 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. 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 a criminal enterprise.

This problem crops up in other settings. Police often have a good idea of who are the local mischief-makers. 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 are directly involved. RICO tried to address this problem, but proving the two separate crimes which are necessary for applying RICO can also be too high a hurdle. Other laws, such as Chicago's gang loitering law,⁸⁵ attempted to address this problem indirectly by permitting the police to disperse groups of people if a known gang member was present, and arrest anyone who failed to comply with orders to disperse. We suspect that this law enabled police to, in effect, aggregate claims against known mischief-makers—people who had committed minor offenses (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

⁸³ In *United States v. Shonubi*, 802 F. Supp. 859 (E.D.N.Y. 1992), 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 intent to distribute it. The court found that in addition to the last occasion where 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 and convicted the defendant for drug possession in the 36th degree, instead of the 28th, thus doubling his prison sentence. The court decision was reversed in *United States v. Shonubi*, 103 F.3d 1085 (2d. Cir.1997).

⁸⁴ 18 U.S.C. §§ 1961–1968 (2006).

⁸⁵ Struck down by the Supreme Court in *Chicago v. Morales*, 527 U.S. 41 (1999). Chicago passed a new anti-gang ordinance after the Supreme Court decided *Morales*. See <http://www.nytimes.com/2000/02/22/us/chicago-council-tries-anew-with-anti-gang-ordinance.html>.

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, is unlikely to be considered suitable for criminal cases. In criminal trials the prosecution, in order to succeed, should prove the defendant's guilt beyond a reasonable doubt, and statistical evidence cannot be the main evidence for conviction.⁸⁶

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 have 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 harm 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 beyond a reasonable doubt, but we can be confident that all the gang members committed a crime against at least some of the victims.

This logic is most clearly visible in the doctrine of joint criminal enterprise liability in international criminal law. The doctrine was invented by the International Tribunal for the Former Yugoslavia in the *Tadic* case.⁸⁷ Tadic belonged to a group of Serbian paramilitaries who engaged in ethnic cleansing in Bosnia. He and his group entered a town for the purpose of intimidating the local Bosnian residents; when they left, five of those residents were dead. Tadic was initially acquitted of murder (which was, in effect, the predicate offense of a crime against humanity, over which the court had jurisdiction), because insufficient evidence connected him to the killings. The Appeals Chamber reversed, holding that Tadic could be convicted because he was part of a joint criminal enterprise and that the killing of the victims was a foreseeable result of that enterprise.⁸⁸ The

⁸⁶ ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 183-5 (2005) (explaining why statistical evidence, standing alone, cannot be the basis for conviction).

⁸⁷ Prosecutor v. Tadic, Case No. IT-94-1-T, Opinion & Judgment (Int'l Crim. Trib. for the Former Yugoslavia May 7, 1997).

⁸⁸ For a lucid discussion, see Allison Marston Danner & Jenny S. Martinez, *Guilty Associations: Joint Criminal Enterprise, Command Responsibility, and the Development of International Criminal Law*, 93 CAL. L. REV. 75, 104-07 (2005).

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 heavily criticized by criminal law 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.⁸⁹ 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.⁹⁰ Consider the following example of cross-claim factual aggregation.

Example IV.1. Targeted Killing: Alternative Claims. The president seeks to use military force to kill a terrorist suspect in Pakistan. However, there is uncertainty both about whether the person is planning an attack on U.S. targets, and about whether the person is a member of Al Qaida. 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 is 64 percent, and thus the

⁸⁹ See, e.g., Danner & Martinez, *supra* note 88; George P. Fletcher & Jens David Ohlin, *Reclaiming Fundamental Principles of Criminal Law in the Darfur Case*, 3 J. INTER'L CRIM. JUSTICE 539 (2005).

⁹⁰ As in the case of tort law, within-claim factual aggregation is ubiquitous and unproblematic. See, e.g., *Salahi v. Obama*, 625 F.3d 745, 753 (D.C. Cir. 2010) (directing the district court to evaluate the habeas petition by considering evidence "collectively rather than in isolation").

correct legal advice would be the opposite.

Now consider a case of cross-element factual aggregation.

Example IV.2. Targeted Killing: Cumulative Claims. 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 percent, and the probability that the target is planning an attack is 60 percent.

We believe that the president's lawyers would advise him that he cannot order the killing of the target because the probability that he is not an American and is planning an attack is only 36 percent (at least if the probabilities are independent—and they may not be). If we are right, then factual aggregation is applied inconsistently—barred in the first case, required in the second case. Our minimal suggestion is that factual aggregation 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.

B. Normative Aggregation

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

Example IV.3. Targeted Killing: Two "Almost" Valid Claims. The president seeks to use military force to kill an American citizen who is alleged to be associated with Al Qaida and who lives in Sana'a, the capital of Yemen. The president claims two sources of authority: a statute that gives him the authority to use military force against Al Qaida; and his constitutional power to use military force abroad to protect American interests. Each claim is at best controversial—not everyone agrees that the statute authorizes actions outside a conventional battlefield, nor that the president can use his constitutional powers to kill an American citizen.

We suspect that most commentators believe that the president may use military force only if each of his claims standing alone can be established. Thus, in Example IV.3, the president would not be permitted to order a targeted killing. Under the aggregation approach, one would reason differently. The president has two "almost" claims—that under the statute he can use force against a terrorist on 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. If we aggregate 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 is fairly common in constitutional adjudication involving the authority of the executive. In *Dames & Moore v. Regan*,⁹¹ for example, the Supreme Court affirmed the president's authority to suspend American claims against Iran, based on aggregation of statutory and constitutional powers:

Although we have declined to conclude that the 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 may be considered to "invite" "measures on independent presidential responsibility."⁹²

The two statutes, the IEEPA and the Hostage Act, do not independently authorize the president's action, but they almost do; and this, 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 provides 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 city. The church argues that the zoning ordinance violates both its constitutional right to free exercise of religion and its constitutional right to 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*, 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

⁹¹ 453 U.S. 654 (1981).

⁹² 453 U.S. at 678.

subject to strict scrutiny,⁹³ 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. 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.⁹⁴

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.”⁹⁵ 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.

The lower courts have heard numerous hybrid cases.⁹⁶ 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 itself. 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

⁹³ *Sherbert v. Verner*, 374 U.S. 398 (1963).

⁹⁴ 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. The Court used the hybrid exception to distinguish *Wisconsin v. Yoder*, where the Court struck down a neutral law because of the burden it imposed on religious association. 406 U.S. 205 (1972). The Court’s reasoning has been harshly criticized by numerous commentators; among other reasons, *Yoder* itself did not 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, 1122 (1990).

⁹⁵ 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).

⁹⁶ 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 either lose or prevail only when the independent claims would independently prevail. The literature on hybrid rights is quite negative; see, e.g., Kyle Still, *Smith’s Hybrid Rights Doctrine and the Pierce Right: An Unintelligent Design*, 85 N.C. L. REV. 385 (2006); 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 94. The authors argue that the doctrine makes no sense on its own terms, and has sowed 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.A. § 2000bb et seq., which we ignore because of our focus on constitutional issues.

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 other threshold of plausibility.⁹⁷

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, albeit 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.⁹⁸

Outside of free exercise, it is difficult to find clear examples of recognition of hybrid rights, but in a number of cases the doctrinal logic suggests such a theory. In *Boy Scouts of America v. Dale*, the Court struck down a state law that forced the Boy Scouts to admit a gay counselor in violation of that organization’s bylaws.⁹⁹ The Court held that the statute violated Boy Scouts members’ right to “expressive association,” which might be taken as a hybrid of the right to free speech and the right to association.¹⁰⁰ In *Griswold v. Connecticut*, the Court struck down a statute that prohibited the sale of contraceptives.¹⁰¹ 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 in the First, Third, and Fourth Amendments.¹⁰² 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*, the Court derived a right of intimate association (such as noninterference in family life) from the right to association in the First Amendment and the right to due process in the Fourteenth Amendment.¹⁰³ Finally, it has been argued that several recent

⁹⁷ See, e.g., *Cornerstone Bible Church v. City of Hastings*, 948 F.2d 464 (8th Cir. 1991) (holding that a church’s challenge to a zoning ordinance was a hybrid claim entitled to strict scrutiny).

⁹⁸ 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 home-schooling, violated free exercise and parental control rights).

⁹⁹ 530 U.S. 640 (2000).

¹⁰⁰ *Id.* at 641. 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,” *id.* at 696, which raised questions as to whether the constitutional standard based only on violations of speech rights should be applied.

¹⁰¹ 381 U.S. 479 (1965).

¹⁰² *Id.* at 484.

¹⁰³ 468 U.S. 609, 618–19 (1984), at least as interpreted by the court in *Marcum v. Catron*, 70 F. Supp. 2d 728, 733–34 (E.D. Ky. 1999) (calling the right to intimate association a “hybrid

Supreme Court cases are best understood as reflecting a hybrid claim involving due process and equal protection rights.¹⁰⁴ In *Lawrence v. Texas*,¹⁰⁵ 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.¹⁰⁶

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—somewhat analogous to the way the unconscionability doctrine was developed so as to aggregate earlier-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*,¹⁰⁷ for example, the owner of a ranch claimed that officials from the Bureau of Land Management (BLM) engaged in a campaign of harassment over a number of years, including trespasses, malicious prosecutions, and the like, in an effort to compel him to grant an easement to the U.S. government. The rancher argued that BLM’s campaign violated the Due Process Clause under the Fifth Amendment, and that BLM retaliated against him for asserting his rights, which was a violation of the First Amendment. The Court acknowledged that the harassment was serious and highly objectionable, but rejected the Fifth Amendment argument because the rancher had various legal remedies that he did not pursue and the conduct fell short of a due process violation; similarly, it rejected the First Amendment claim because the government’s purpose—to obtain land—was not as clearly improper as its purpose in traditional retaliation claims where the government objects to protected speech. What is notable about this case is that the court evidently believed that BLM had acted wrongfully, but did not consider the possibility that even if BLM did not violate the Due Process Clause and the Free Speech Clause individually, it did violate the two of them taken together (or a third, such as the Equal Protection Clause, which prohibits government conduct motivated by animus¹⁰⁸), by analogy to the hybrid rights cases. Indeed, this

right”).

¹⁰⁴ Kenji Yoshino, *The New Equal Protection*, 124 HARV. L. REV. 747 (2011).

¹⁰⁵ 539 U.S. 558 (2003).

¹⁰⁶ Yoshino, *supra* note 104, at 778–79.

¹⁰⁷ 551 U.S. 537 (2007).

¹⁰⁸ *See, e.g., Village of Willowbrook v. Olech*, 528 U.S. 562 (2000).

approach is the norm—the hybrids 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*, a case in which a defendant was sentenced to a term of 37 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.¹⁰⁹ The defendant argued that reimprisonment after such a delay violated his rights to substantive and procedural due process. 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 Court held that Sanders' sentence had been appealed by both sides, so Sanders had no reason to believe that his sentence was final. The Court also rejected Sanders' 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 tactical advantage. The substantive due process and procedural due process claims were 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."¹¹⁰ The dissent, in essence, argued that even if the substantive due process violation was not as serious as in other cases (Sanders' 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 entitled Sanders to relief. 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 that were 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 1.7, which involved a mass tort with indeterminate victims. Defendant's factory pollutes, creating a statistical likelihood of

¹⁰⁹ 432 F.3d 572 (6th Cir. 2006).

¹¹⁰ *Id.* at 583.

harming an additional 25 people per year, but none of those people can be identified. Tort law does not permit cross-person aggregation; a remedy is not available. 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.¹¹¹ 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 (EPA) 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 receives a remedy, but in future there will be fewer victims. This is a clear example of cross-person aggregation.¹¹²

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 examples,¹¹³ a prominent example involves religious displays on public property, which are forbidden under the First Amendment if they are sectarian.¹¹⁴ One way of thinking about religious displays is that they inflict

¹¹¹ 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).

¹¹² 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, but they can collectively.

¹¹³ See, e.g., *Massachusetts v. EPA*, 549 U.S. 497 (2007) (holding that state has standing to challenge 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 state has standing to bring tort claim against polluter on behalf of its citizens).

¹¹⁴ See *Lemon v. Kurtzman*, 403 U.S. 602 (1971); *County of Allegheny v. American Civil*

a non-physical, and hence difficult-to-prove but nonetheless important psychic harm on people outside the religion that the display celebrates. Accordingly, 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, it is statistically likely that at least one of them was harmed. The law gives even a single individual standing to bring a claim, but in effect on behalf of the group. This approach can be contrasted to the law's reluctance to give remedies to tort victims who claim emotional but not physical harm.¹¹⁵ 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.¹¹⁶

V. EXPLANATIONS AND PROPOSALS

A. *The Arbitrariness of Legal Boundaries*

The best explanation for the 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 effect of tort law is twofold: to optimally deter people from imposing externalities on each other and sometimes¹¹⁷ to compensate people who have suffered from those externalities. 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 be required to aggregate harms and probabilities in all circumstances. So, for example, the decisionmaker would hold liable the defendant who acted negligently with 40 percent probability while driving a vehicle that was inherently dangerous with 40 percent probability (Example I.1). The explanation is straightforward. 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, she will engage in those actions,

Liberties Union Greater Pittsburgh Chapter, 492 U.S. 573 (1989).

¹¹⁵ EPSTEIN, *supra* note 27, at 274-75 (noting that recovery for emotional harm without physical impact is generally denied everywhere today).

¹¹⁶ But see *supra* text accompanying notes 41-44.

¹¹⁷ Under a negligence rule, only if the externalities were inefficiently (or unreasonably) imposed.

even though in an expected sense they will cause harm which is higher than their benefit. By holding such a person liable, the ideal decisionmaker deters socially costly behavior.

But as the literature on rules versus standards makes clear,¹¹⁸ such a system of pure standards would not function effectively in a world in which decision costs are often high. The preponderance of the evidence rule 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.¹¹⁹

We turn now to normative aggregation. Take Example II.2 (Non-Material Breach and Minor Fraudulent Misrepresentation), where defendant engages in a minor form of fraudulent misrepresentation in order to secure 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 fraudulent misrepresentation crosses a threshold, and (2) whether the breach crosses a threshold, the court would be required to determine (1) whether the fraudulent misrepresentation crosses a threshold, (2) whether the breach crosses a threshold, *and* (3) whether the *combined* actions cross a *different* threshold. This more complicated test may well create an unacceptable level of difficulty and uncertainty. But non-aggregation 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-claim 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 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 limitation 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.¹²⁰ By contrast, when the claims are of the same

¹¹⁸ *E.g.*, Louis Kaplow, *Rules versus Standards: An Economic Analysis*, 42 DUKE L. J. 557 (1993).

¹¹⁹ For the difficulties of factual aggregation, *see* Levmore, *supra* note 7, at 726–33. Most of Levmore's examples come from tort law, and relate to what we called cross-element aggregation.

¹²⁰ Note that these questions are much less acute when the plaintiff is able to establish both

nature, no such conflicts arise.

Should the law aggregate more than it does? Our minimal suggestion is that the doctrine is currently inconsistent, but we suspect that the answer is yes. Courts should recognize when they aggregate and when they do not, and explain why they aggregate in some cases and not others. They should be open to aggregation arguments from litigants, which in many cases could change outcomes. They should also realize that litigants who make aggregation arguments can draw analogies to established norms and patterns in the law that already recognize aggregation, as we have shown. And when decision costs are low, factual aggregation and normative aggregation should clearly improve legal outcomes. 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. To be sure, there are cases in which aggregation is unnecessary and possibly harmful, and we will address them in the Conclusion.

B. Other Explanations 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 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.¹²¹ Moreover, under corrective justice theories, 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.¹²²

tort and contract claims separately. In the latter case the plaintiff would generally be entitled to the remedy that is more favorable to him.

¹²¹ JULES COLEMAN, *RISK AND WRONGS* 380–85 (2003) (discussing the focus of corrective justice on the relationship between the injurer and the victim); ERNEST WEINRIB, *THE IDEA OF PRIVATE LAW* 63–66 (1995) (same).

¹²² 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 121, 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.”)

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.7 (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.¹²³ Market share liability, however, could be reconciled at least with some versions of corrective justice.¹²⁴

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, none 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 one another. Thus, in Example I.3 (Injury in the Hospital: Two Events, Two Injuries), a patient suffered two distinct harms, each of which might be caused by a different doctor. The hospital is vicariously liable for both doctors' wrongdoing. While the patient cannot establish the liability of each doctor, 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

¹²³ WEINRIB, *supra* note 121, 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); Stephen R. Perry, *Risk, Harm, and Responsibility*, in *PHILOSOPHICAL FOUNDATIONS OF TORT 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).

¹²⁴ 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: a group of wrongdoers inflict 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 20, at 132–33.

determination of liability should be done per event, and not across events (we suspect that corrective justice theorists would persist on that view, even if in our example the two harms might have been the result of the same doctor's wrongdoing).¹²⁵ 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 views.¹²⁶ A possible explanation for courts' reluctance to engage in some aggregations but not in others is their refusal to evaluate claims of a different nature according to one common scale.

Take examples II.1 (Either Material Breach or Fraudulent Misrepresentation), and II.2 (Non-Material Breach and Minor Fraudulent Misrepresentation). Both examples deal with a case where the plaintiff argues that he was entitled to rescind a contract since the defendant engaged in fraudulent misrepresentation in order to secure plaintiff's consent to the contract and subsequently 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, both the misrepresentation and the breach, standing alone, are not severe enough to justify the rescission of the contract. In order to aggregate both 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 just 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 Example the

¹²⁵ Corrective justice may support liability if the two events can be reasonably understood as one event occurring in two stages.

¹²⁶ See generally, INCOMMENSURABILITY, INCOMPARABILITY, AND PRACTICAL REASON (Ruth Chang ed., 1997). See also 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); see also 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); Amartya Sen, *Plural Utility*, 81 PROC. ARISTOTELIAN SOC’Y 193, 193–210 (1981) (arguing that welfare economics should understand utility “primarily as a vector (with several distinct components), and only secondarily as some homogeneous magnitude”).

court would have to decide whether “almost” fraudulent misrepresentation combined with “almost” material breach, are 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 much more than aggregating probabilities in the other examples which 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 constitutional law example IV.3 (Targeted Killing: Cumulative 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 Qaida, and his constitutional power to use military force abroad to protect American interests. Here too, assuming neither of the legal sources provides the authority for target 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. Indeed, sometimes comparing values of a different nature raises implementation difficulties, and those difficulties should be taken into account in considering the desirability of some kinds of aggregations. But then the commensurability argument appears to be just another way to make the argument that aggregation has costs of implementation, as we have shown in section 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 plaintiff if either

claim #1 or claim #2 is supported by preponderance of the evidence, or if claim #1 and claim #2 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 in likelihood, 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.¹²⁷

A similar problem could also upset normative aggregation. For Example II.1 (Either Material Breach or Fraudulent Misrepresentation), the court would instruct the jury to “find for plaintiff if defendant’s statements were fraudulent, 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 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 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.¹²⁸

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 not because they might have collaborated or drawn on the same source of knowledge. Thus, while cognitive limitations should play a role in the design of legal doctrine, including the uses of aggregation, they cannot by themselves provide a sufficient reason for rejecting aggregation.

¹²⁷ 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 PSYCHOLOGY, PUB. POL’Y, & LAW 769 (2000) (discussing psychological literature on jurors’ ability to distinguish standards of proof). Levmore, *supra* note 7, at 739–45, makes the interesting argument, based on the Condorcet Jury Theorem, that if all (or a supermajority of the) members of the jury believes that a factual allegation is more probable than not, say, by a probability of 51%, than the probability of the allegation being true is in fact much higher than 51%. Levmore suggests that this could be a reason not to aggregate, mainly in cases that we called cross-element aggregation in this Article.

¹²⁸ The research described in Clermont, *supra* note 127, bears on this question as well.

CONCLUSION

We have analyzed three types of aggregation in the law through various examples in four central legal fields. In most of our examples, courts do 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-claim 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 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 aggregation of probabilities cancel out ex post, and so have no effect on ex ante incentives, where the law is concerned only with those ex ante incentives.¹²⁹ 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.¹³⁰

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 accurately; 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 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-claim aggregation. Instead of asking whether claim A is more probable than not, they should 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

¹²⁹ See *supra* text accompanying notes 18–21.

¹³⁰ See *supra* text accompanying note 25.

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 is even more socially desirable rather than less.¹³¹ 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.¹³²

It may also be difficult for decisionmakers to make 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 a 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 decision-making has an ad hoc, almost remedial quality. The plaintiff shows that she has two claims, each has normative weight which 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 for constructing new claims. The unconscionability doctrine illustrates this process. A plaintiff has two claims under an old legal regime which are individually weak but jointly powerful; rather than give 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.¹³³

¹³¹ See *supra* text following note 29.

¹³² EYAL ZAMIR AND BARAK MEDINA, LAW, ECONOMICS, AND MORALITY 46 (2010) (“...moderate deontology holds that constraints have thresholds. A constraint may be overridden for the sake of furthering good outcomes or avoiding bad ones if enough good or bad is at stake”).

¹³³ It has been frequently argued that the law has been evolving from rules toward standards.

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 when, however, the identity of her specific victims cannot be ascertained. 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 life they tend *not* to expose this way of reasoning to other people. It would not be surprising if judges and jurors also aggregate more than they say they do.

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 described in the text, 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.