AN OBJECTIVE-CHANCE EXCEPTION TO THE RULE AGAINST CHARACTER EVIDENCE

Hillel J. Bavli

AN OBJECTIVE-CHANCE EXCEPTION TO THE RULE AGAINST CHARACTER EVIDENCE

*Hillel J. Bavli**

A central principle of U.S. law is that individuals should be judged in court based on their actions and not on their character. Rule 404 of the Federal Rules of Evidence therefore prohibits evidence of an individual's previous acts to prove that the individual acted in accordance with a certain character or propensity. But, courts regularly deviate from or altogether ignore this rule, resulting in arbitrariness and judgments based on an individual's prior acts rather than on evidence regarding the events at issue in a case.

In this Article, I argue that at the center of the unpredictability surrounding the rule against character evidence is a type of evidence that I refer to as "objective-chance evidence"—that is, evidence regarding other events of the same general kind as the event in question, offered to show that the event in question is due to some intent or design rather than to accident or chance. I apply simple scientific principles of information aggregation to examine the nature of objective-chance evidence in the courts and literature. I then argue that central to a more logical and effective approach to character evidence are (1) a proper understanding of objective-chance evidence as a particular category of character evidence, and (2) an "objective-chance exception" that replaces the rule against character evidence with a Rule 403 balancing for objective-chance evidence. I show that these conditions may permit a more coherent interpretation of Rule 404 and ultimately a stricter adherence to the rule against character evidence.

INTRODUCTION

The recent *Chauvin* trial, *Rittenhouse* trial, and *Cosby* trial all involved critical decisions regarding whether the court should admit evidence regarding prior acts committed by the defendant or a victim.[1] This evidentiary decision appears in a large proportion of criminal cases and is correctly understood as highly impactful on a case's outcome.[2] Indeed, we regularly make predictions and infer

* Assistant Professor of Law, SMU Dedman School of Law. For helpful comments and exchanges, I thank William Hubbard, Barry Goldstein, Edward Cheng, Paul Rothstein, Edward Imwinkelried, Teneille Brown, Jenia Turner, Daniel Capra, Liesa Richter, George Fisher, Timothy Lau, Frederick Moss, and the participants of the 2021 Evidence Summer Workshop at Vanderbilt Law School and the Fall 2021 SMU Dedman School of Law Faculty Workshop. I also thank Laura Sundin and Hyewon Choi for their outstanding research assistance. I am grateful for generous financial support provided by SMU Dedman School of Law and the Glenn Portman Faculty Research Fund during the completion of this Article in the summer of 2020.

1. *See* Commonwealth v. Cosby, 252 A.3d 1092, 1121–26 (Pa. 2021); State v. Chauvin, No. 27-CR-20-12646, 2021 WL 252713, at *4 (D. Minn. Jan. 26, 2021); NBC Chicago, *Full Video: Prosecutors Cross-Examine Kyle Rittenhouse*, YOUTUBE (Nov. 11, 2021, 32:00), https://www.youtube.com/watch?v=JG8PhtFrO0Y&t=1964s.

2. *See* United States v. Burkhart, 458 F.2d 201, 204 (10th Cir. 1972) ("[A]n obvious truth is that once prior convictions are introduced the trial is, for all practical purposes, completed and the guilty outcome follows as a mere formality. This is true regardless of the care and caution employed by the court in instructing the jury."); Paul S. Milich, *The Degrading Character Rule in American Criminal Trials*, 47 GA. L. REV. 775, 780 (2013) ("The stakes for the prosecution and the defense are enormous. Once the jury learns that the defendant has a criminal past, the odds of conviction skyrocket."); *see also* United States v. Davis, 726 F.3d 434, 441 (3d Cir. 2013) ("Rule 404(b) has become the most cited evidentiary rule on appeal."); Jeffrey Bellin, *The Evidence Rules That Convict the Innocent*, 106 CORNELL L. REV. 305, 343 (2021) ("Some of the most heated controversies in the modern evidence landscape involve evidence of uncharged crimes admitted under Federal Rule of Evidence 404(b) or innovative variants like Rules 413 and 414."); Edward J. Imwinkelried, *The Use of Evidence*

facts based on our knowledge of prior behavior. Employers request references before hiring an employee; homeowners seek recommendations before hiring a contractor; and we tend to assume that the culprit of a violent crime is more likely to be an individual who has committed similar violent acts in the past than an individual who has not. Under certain conditions, these inferences are central to good decision-making; under others, they can lead to incorrect judgments and severe consequences.

The concern of this Article is the question, when should courts permit a factfinder to make such inferences, and when should courts forbid them? The difference between admissibility and inadmissibility can often change the outcome of a case. It can often mean the difference between a verdict of guilty and a verdict of not guilty.

The Federal Rules of Evidence (FRE) generally forbid character-based propensity reasoning.[3] In particular, under Rule 404, a party may not offer evidence of a person's character—either directly or via the person's prior behavior—to demonstrate "that on a particular occasion the person acted in accordance with the character."[4] This rule reflects the legal principle that individuals should be judged based on what they did and not based on who they are.[5] Under Rule 403, a court would apply a balancing test that compares the probative value of character evidence with the unfair prejudice that would result from it.[6] But Rule 404 preempts this balancing and replaces it with a general rule against character evidence.[7]

Rule 404(b)—the provision of the character evidence rule that pertains to "[e]vidence of any other crime, wrong, or act"—is among the most cited, and arguably the most controversial, of the FRE.[8] This is for good reason: other-acts character evidence can be both very probative and very prejudicial. The highly prejudicial nature of this evidence underlies the wholesale ban against it.[9] At the same time, however, courts frequently admit other-acts character evidence.[10] Courts and legislatures have chipped away at the prohibition against character evidence as they have yielded to pressure created by the substantial probative value that other-acts character evidence frequently entails. Some

---

of an Accused's Uncharged Misconduct to Prove Mens Rea*: The Doctrines Which Threaten to Engulf the Character Evidence Prohibition*, 51 OHIO STATE L.J. 575, 577 (1990) ("Rule 404(b) has generated more published opinions than any other subsection of the Federal Rules.").

3. *See* FED. R. EVID. 404(a), (b).

4. FED. R. EVID. 404(a)(1), (b)(1).

5. *See* United States v. Gomez, 763 F.3d 845, 861 (7th Cir. 2014); United States v. Brown, 597 F.3d 399, 404 (D.C. Cir. 2010).

6. *See* FED. R. EVID. 403.

7. Throughout this Article, although I refer to the Federal Rules of Evidence, my analysis generally applies to state law also.

8. FED. R. EVID. 404(b)(1); *see* Edward J. Imwinkelried, *A Brief Essay Defending the Doctrine of Objective Chances as a Valid Theory for Introducing Evidence of an Accused's Uncharged Misconduct*, 50 N.M. L. REV. 1, 1 (2020).

9. *See infra* Subpart I.A.

10. *See infra* Subparts I.B–C.

exceptions to Rule 404, such as those created to allow character-based propensity reasoning in sexual assault cases, have been established formally through legislation.[11] Others have surfaced through common law.[12] Frequently, however, courts have permitted character evidence informally—often through misreadings of Rule 404(b)—or they have ignored it altogether.[13]

For example, Rule 404(b)(2), which provides for permitted *non-propensity* uses of other-acts evidence—such as to show knowledge, motive, or absence of mistake or accident—is often misinterpreted as an *exception* to Rule 404(b)(1)'s rule against other-acts character evidence rather than a mere clarification that emphasizes the permissibility of other-acts evidence that does not rely on propensity reasoning.[14] For example, many courts admit evidence regarding a defendant's prior drug-related crimes to prove that the defendant had knowledge or intent related to a drug crime in question.[15]

The ban on character evidence is so riddled with exceptions, inconsistencies, complexities, and ad hoc deviations that it cannot be said to exclude impermissible character evidence with any predictability, except perhaps in its most blatant form.[16] While Rule 404 exists to create consistency and predictability with respect to the admissibility of character evidence, it has not only failed to achieve this goal, but, in some contexts, it has arguably added confusion and uncertainty. Consequently, the admissibility or inadmissibility of other-acts character evidence—a category of evidence that is undoubtedly highly impactful on the outcome of a case—is frequently left to chance. This leads to arbitrary outcomes and verdicts based on acts not at issue in a case.

In a previous Article, I developed a model to examine the effect of character evidence on accuracy and to explain why and when courts depart from the rule against character evidence.[17] I applied this "aggregation-evidence" model to demonstrate the extraordinary nature of a particular type of evidence that I refer to as "objective-chance evidence," in terms of its ability to improve accuracy.[18] This evidence involves information regarding other events of the same general kind as the event at issue in a case, offered to show that the event at issue is unlikely to be due to accident or chance.[19] For example, in a case in which a man argued that his wife's bathtub drowning was an accident, evidence

11. *See* FED. R. EVID. 413–15.

12. *See infra* Subpart I.B.

13. *See infra* Subparts I.B–C.

14. FED. R. EVID. 404(b)(2).

15. *See infra* notes 65–69 and accompanying text.

16. *See infra* Subparts I.B–C.

17. *See* Hillel J. Bavli, *An Aggregation Theory of Character Evidence*, 51 J. LEGAL STUD. 39 (2022) [hereinafter *An Aggregation Theory of Character Evidence*].

18. *Id.* at 54–58.

19. *See id.* at 41; Imwinkelried, *supra* note 8, at 4–9 (explaining the doctrine of objective chances and distinguishing it from character reasoning); GEORGE FISHER, EVIDENCE 193–201 (3d ed. 2013) (discussing the doctrine of objective chances); *see also infra* Subpart I.D.

that two subsequent wives of the defendant also drowned in their baths was permitted for the inference that the drowning was not an accident but rather was due to the design of the defendant.[20]

In that Article, I argued that objective-chance evidence is central to many of the courts' departures from the rule against character evidence.[21] In particular, the tendency of objective-chance evidence to substantially improve accuracy creates pressure on courts to admit this type of evidence, notwithstanding the general ban on character evidence.[22] Consequently, courts find ways around the rule against character evidence through exceptions or loose reasoning. These decisions are then propagated in future cases, resulting in further misapplications of Rule 404(b) and broader uncertainty surrounding the rule.

In the current Article, I argue that to address the state of disarray surrounding the rule against character evidence, it may be necessary to establish an "objective-chance exception" that would replace the rule against character evidence with a Rule 403 balancing analysis for objective-chance evidence. I develop my argument in three stages. First, I argue that the problem of judicial deviations from the rule against character evidence is even greater than is currently recognized. Indeed, in some areas of the law, the admission of certain types of character evidence is so deeply ingrained in the common law that courts simply cite to generic precedent and thereby altogether bypass any explanation for their admission of character evidence with respect to Rule 404.[23]

Second, I apply the aggregation-evidence model to argue that a significant source of the disarray surrounding the rule against character evidence is a misclassification of objective-chance evidence and the "doctrine of objective chances" (or the "doctrine of chances")[24] as involving evidence not requiring propensity reasoning—that is, as involving evidence separate and distinct from character evidence.[25] Specifically, I apply an area of statistics known as Bayesian inference—which essentially provides a set of rules for combining prior beliefs with new evidence to arrive at a new belief or conclusion—to unravel the true nature of objective-chance evidence and to demonstrate that objective-chance

---

20. *See* Rex v. Smith, (1915) 11 Cr. App. R. 229, 84 LJKB 2153; *see also* United States v. Henthorn, 864 F.3d 1241 (10th Cir. 2017).

21. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 40–41, 58–61.

22. *Id.*

23. *See infra* Subpart I.C.

24. Imwinkelried, *supra* note 8, at 6–7.

25. *See infra* Parts I, III; *see, e.g.*, *Smith*, 84 LJKB at 2153–54; Sean P. Sullivan, *Probative Inference from Phenomenal Coincidence: Demystifying the Doctrine of Chances*, 14 LAW, PROBABILITY & RISK 27 (2015); Edward J. Imwinkelried, *An Evidentiary Paradox: Defending the Character Evidence Prohibition by Upholding a Non-Character Theory of Logical Relevance, The Doctrine of Chances*, 40 U. RICH. L. REV. 419 (2006) [hereinafter Imwinkelried, *An Evidentiary Paradox*]; Imwinkelried, *supra* note 2, at 585–601. *But see* Paul F. Rothstein, *Comment*: *The Doctrine of Chances, Brides of the Bath and a Reply to Sean Sullivan*, 14 LAW, PROBABILITY & RISK 51 (2015); Paul F. Rothstein, *Intellectual Coherence in an Evidence Code*, 28 LOY. L.A. L. REV. 1259 (1995) [hereinafter Rothstein, *Intellectual Coherence in an Evidence Code*].

evidence is best understood (even if not always required to be understood) as a form of character evidence.[26]

Lastly, however, I argue that in light of the unique nature of objective-chance evidence—and specifically, its extraordinary ability to improve accuracy—a more coherent rule against character evidence may require an objective-chance exception that replaces the rule against character evidence with a Rule 403 balancing for objective-chance evidence. I argue that such an exception would allow courts to admit this uniquely probative category of character evidence without misclassifying it and without causing uncertainty and poor law for future cases. At the same time, it would promote an otherwise stricter adherence to Rule 404, including better admissibility decisions under Rule 404(b)(2)'s provision regarding non-character uses of other-acts evidence.

In particular, courts frequently admit objective-chance evidence under one of two false premises: (1) the misidentification of the evidence as non-character evidence, or (2) the misapplication of Rule 404(b)(2) to admit other-acts evidence even when it relies on propensity reasoning.[27] In turn, these admissibility decisions result in misapplications of the law and the propagation of incorrect law for future cases—including misinterpretations of Rule 404(b)(2), incorrect applications of the doctrine of chances, and ad hoc carve-outs from the ban on character evidence.[28] Courts then further promote such deviations from Rule 404 by simply relying on this precedent rather than addressing Rule 404 explicitly. For example, a court may admit objective-chance evidence by misinterpreting Rule 404(b)(2) as providing for a broad range of exceptions to the rule against character evidence. A future court may then cite this precedent even for non-objective-chance evidence—that is, for ordinary character evidence—as a basis for admissibility.[29] An objective-chance exception would address this problem at its source.

Courts tend to depart from Rule 404 in circumstances in which adhering to it means making counterintuitive or unreasonable evidentiary exclusions. The substantial accuracy benefits associated with objective-chance evidence create pressure for courts to depart from Rule 404 in a wide range of contexts. By formulating a principled and well-defined objective-chance exception, the proposed approach seeks to alleviate this pressure and permit a sensible and proper interpretation of Rule 404(b)(2) and a stricter adherence to the rule against character evidence. The proposed method thereby aims to eliminate the tendency toward haphazard and unpredictable admissibility decisions under Rule 404 while permitting admissibility for a category of character evidence that stands apart in its ability to improve accuracy.

---

26. *See infra* Parts II, III. *See generally* PETER D. HOFF, A FIRST COURSE IN BAYESIAN STATISTICAL METHODS 1 (2009) (discussing Bayesian inference).

27. *See infra* Subparts I.B, III.B.

28. *See infra* notes 68–69 and accompanying text.

29. *See infra* Subparts I.B–C.

I proceed as follows: In Part I, I examine the law and critical policy objectives surrounding the rule against character evidence. I then discuss the courts' routine departures from the rule against character evidence, and I argue that the problem is even more severe than is commonly recognized. In Part II, I describe the aggregation-evidence model and its implications for the effect of character evidence on accuracy and for the extraordinary nature of objective-chance evidence. In Part III, I apply the aggregation-evidence model to argue that objective-chance evidence should be understood and treated as a particular category of character evidence. I then apply my analysis to develop an objective-chance exception to the rule against character evidence and to argue that such an exception may lead to a more logical and effective rule.

I. THE RULE AGAINST CHARACTER EVIDENCE

Rule 404 of the FRE codifies the common-law rule that evidence of an individual's character may not be introduced to prove that the individual acted in accordance with that character on a particular occasion.[30] It reflects a "foundational principle in our system of justice that 'we try cases, rather than persons'"—that individuals should be judged based on what they did rather than on who they are.[31] Subsection 404(a)(1) provides, "Evidence of a person's character or character trait is not admissible to prove that on a particular occasion the person acted in accordance with the character or trait."[32] Subsection 404(b)(1) addresses evidence of a person's character via their prior actions, providing that "[e]vidence of any other crime, wrong, or act is not admissible to prove a person's character in order to show that on a particular occasion the person acted in accordance with the character."[33]

In this Part, I begin by examining the policy rationales underlying Rule 404. I then discuss the state of disarray that surrounds the courts' application of Rule 404, and I discuss scholarship surrounding the admissibility of other-acts evidence under the doctrine of objective chances.

A. Rule 404's Purpose: Accuracy and the "Action-Not-Person" Principle

The primary purpose of Rule 404 is to avoid two forms of unfair prejudice:[34] (1) the risk that the factfinder will give excessive weight to character evidence in determining a verdict, and (2) the risk that the factfinder will punish

30. *See* FED. R. EVID. 404; FISHER, *supra* note 19, at 153.

31. United States v. Gomez, 763 F.3d 845, 861 (7th Cir. 2014) (quoting People v. Allen, 420 N.W.2d 499, 504 (Mich. 1988)); *accord* United States v. Brown, 597 F.3d 399, 404 (D.C. Cir. 2010).

32. FED. R. EVID. 404(a)(1).

33. FED. R. EVID. 404(b)(1).

34. Throughout this Article, although I sometimes use the term "unfair prejudice" for emphasis, I generally use the term "prejudice" to entail implicit unfairness.

the party against whom the evidence is offered based, at least in part, on their character or prior acts rather than on the events at issue in a case.[35] For example, if the prosecutor in a drug-trafficking case is permitted to introduce evidence that the defendant has been arrested for drug trafficking five times over the past five years, a jury may afford this evidence excessive weight in determining whether the defendant committed the drug-trafficking crime with which they are presently charged. Additionally, the jury may hold the prosecution to a lower standard of proof—or, worse, simply convict the defendant—on the basis of the defendant's previous arrests in particular.

There is some controversy and confusion regarding whether the primary aims of Rule 404 are grounded in accuracy or in other policy objectives. This is important. To start, the primary aim of evidence law, and the FRE in particular, is accuracy. "We want juries to return the right verdict, and by that we may mean the *truthful* verdict, the one that accords with *what happened*."[36] Rule 102 of the FRE states: "These rules should be construed so as to administer every proceeding fairly, eliminate unjustifiable expense and delay, and promote the development of evidence law, to the end of ascertaining the truth and securing a just determination."[37] Next, it is clear from the Rules that they are also influenced by, and sometimes based entirely on, policies other than achieving accuracy. For example, the privilege that protects confidential communications between spouses from admission at trial aims to protect the sanctity of marital communications and promote free and open communication between spouses and good marital relationships—even at the cost of sacrificing accuracy at trial.[38]

What, then, are the primary aims of Rule 404? Preliminarily, it is well-accepted that character evidence can be very probative. This proposition is supported by common experience, empirical research, and court decisions.[39] It is also well-evidenced by the courts' frequent admission of other-acts character evidence, notwithstanding Rule 404 and notwithstanding a balancing of probative value and unfair prejudice under Rule 403, where this type of

35. *See* People v. Zackowitz, 172 N.E. 466, 468 (N.Y. 1930) (citing 1 JOHN HENRY WIGMORE, WIGMORE ON EVIDENCE § 194 (1904)); FISHER, *supra* note 19, at 153. Note that there are other dangers that the introduction of character evidence risks—for example, the danger of causing confusion and of wasting time. *See* FISHER, *supra* note 19, at 154.

36. FISHER, *supra* note 19, at 1.

37. FED. R. EVID. 102.

38. *See* CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, EVIDENCE UNDER THE RULES: TEXT, CASES, AND PROBLEMS 867–68 (8th ed. 2015); FISHER, *supra* note 19, at 1056–57.

39. *See, e.g.*, *Zackowitz*, 172 N.E. at 468 ("The principle back of the exclusion is one, not of logic, but of policy. There may be cogency in the argument that a quarrelsome defendant is more likely to start a quarrel than one of milder type, a man of dangerous mode of life more likely than a shy recluse. The law is not blind to this, but equally it is not blind to the peril to the innocent if character is accepted as probative of crime." (citations omitted)). *But see* FED. R. EVID. 404 advisory committee's note to proposed rule (suggesting that character evidence, at least under subsection (a), entails little probative value).

evidence is, without dispute, highly prejudicial.[40] The rule against character evidence is certainly not rooted in an absence of probative value. Rather, character evidence is excluded from trial based on the risk of undue prejudice that it entails.[41]

Therefore, let us next ask whether the risk of undue prejudice caused by character evidence is rooted in a concern for accuracy or some other policy. For example, when we fear that a jury will afford other-acts character evidence excessive weight, is it because we believe that jurors will tend to think that people act in line with their character more than they actually do, or is it simply because there is a strong legal norm to judge a person based on his actions rather than on his character—in which case, character evidence should be afforded less weight *notwithstanding its accuracy benefits*?

The ban on character evidence likely arises from both accuracy and non-accuracy policy concerns. There is undoubtedly a fear that the jury will be drawn away from the correct outcome in a case. For example, the fear that a jury will seek to punish a defendant for past crimes, regardless of the strength of the evidence proving that the defendant committed the crime at issue, clearly implicates accuracy concerns. Similarly, the fear that a jury will afford more weight to character evidence than is justified by the current understanding of an individual's tendency to act in line with a particular character trait also implicates important accuracy concerns.

On the other hand, the ban on character evidence, at least as expressed formally, does not seem to account for or fluctuate based on our changing understanding of a person's tendency to act in line with their character. It does not seem sensitive to these changes. Moreover, Rule 404 is frequently characterized by the Advisory Committee, courts, and scholars as a particular application of Rule 403 balancing.[42] Rule 403 provides that "[t]he court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence."[43] It is unlikely, however, that the prejudice caused by character evidence with respect to accuracy alone would, as a general matter (and sufficiently general to justify a wholesale exclusion of character evidence), *substantially outweigh* the probative value of such evidence.[44] Only by incorporating a special policy against character-based judgments—in particular, a special aversion to, and a weight against, prejudice caused by character

40. *See infra* Subparts I.B–C.

41. *See supra* note 35 and accompanying text.

42. *See* FISHER, *supra* note 19, at 154 ("Rule 404 reflects the judgment of Congress that *as a matter of law* the probative value of propensity evidence is substantially outweighed by the risk it poses of unfair prejudice, juror confusion, and waste of time.").

43. FED. R. EVID. 403.

44. *Id.*

evidence—could a Rule 403 balancing justify Rule 404's wholesale exclusion of character evidence.[45]

Rule 404 thus entails more than simply a special application of Rule 403. Rather, it is one that places particular weight on the prejudice caused by character evidence. This particular weight reflects the legal principle that we judge individuals based on their actions rather than on their character, even when doing so reduces accuracy on average.[46]

B. The Courts' Unprincipled Approach to Other-Acts Character Evidence

Rule 404(b) contains an important application of Rule 404(a)'s rule against character evidence. It is the focus of this Article, and it is arguably the most controversial rule in the FRE.[47] Rule 404(b)(1) disallows evidence of an individual's "other crime, wrong, or act" to prove character for the purpose of showing that they acted in line with that character on a particular occasion.[48] Other-acts evidence can be extremely probative or extremely prejudicial because it concretely informs the factfinder of other crimes, wrongs, or acts of an individual. The admission of other-acts evidence can often be determinative of a case's outcome.[49]

Sometimes, other-acts evidence is probative of a matter *not* requiring propensity reasoning. For example, a previous assault or attempted murder by the defendant against a victim may be offered to prove that the defendant had a motive to assault or kill the victim on the occasion at issue in a case. A previous act in which a defendant bypassed a particularly complex security system may be offered to show that the defendant had the knowledge or skill to bypass that system on the occasion in question. Such non-propensity uses do not invoke the same concerns as character-based propensity reasoning, and when other-acts evidence is used for non-propensity purposes, it does not constitute character evidence. Therefore, pursuant to Rule 404(b)(2), other-acts evidence "may be admissible for another purpose, such as proving motive,

45. This idea is recognized in Justice Cardozo's decision in *Zackowitz*. Justice Cardozo makes clear that the character evidence in *Zackowitz* may well be probative of the defendant's guilt and even accuracy-enhancing: "The principle back of the exclusion is one, not of logic, but of policy." People v. Zackowitz, 172 N.E. 466, 468 (N.Y. 1930). However, the evidence is too prejudicial in light of other policies: "The law is not blind to this, but equally it is not blind to the peril to the innocent if character is accepted as probative of crime." *Id.* That is, in light of the law's policy interest in protecting against false convictions—a particular application of the principle that judgments should be based on actions and not character, and of the law's aversion to character-based prejudice—the law will adopt an understanding of the evidence as net-prejudicial rather than net-probative.

46. *See supra* note 31 and accompanying text.

47. Imwinkelried, *supra* note 8, at 1.

48. FED. R. EVID. 404(b)(1).

49. *See* Imwinkelried, *supra* note 8, at 1–2.

opportunity, intent, preparation, plan, knowledge, identity, absence of mistake, or lack of accident."[50]

This subsection is a clarification rather than an exception: the rule against character evidence in Rule 404(a)(1) and Rule 404(b)(1) disallows propensity reasoning, but other-acts evidence offered for non-propensity purposes does not violate the rule.[51] Under this reading, Rule 404(b)(2) simply clarifies that if such evidence is offered for non-propensity purposes, it may be admissible. In this case, under Rule 403, the evidence is admissible only if the probative value of the evidence with respect to the permissible purpose is not substantially outweighed by the risk of unfair prejudice associated with the possibility of impermissible propensity reasoning.[52]

Importantly, understanding Rule 404(b)(2) as a clarification of Rule 404(b)(1)—rather than as providing for exceptions to Rule 404(b)(1)—is the only sensible reading of the rule. Understanding Rule 404(b)(2) as providing for exceptions to Rule 404(b)(1) would allow this provision to altogether swallow the rule against other-acts character evidence. After all, it is simple to articulate even the most prejudicial forms of character evidence—the precise type of evidence that is intended to be excluded by Rule 404—as permissible evidence under Rule 404(b)(2) if this provision is understood as an exception rather than a clarification. For example, assume that a prosecutor seeks to introduce evidence that the defendant has committed two previous robberies for the purpose of demonstrating that the defendant has a propensity to commit such crimes and is therefore likely to have acted pursuant to this propensity by committing the act in question. If Rule 404(b)(2) is viewed as an exception, this type of propensity reasoning is not off limits for the prosecutor. Rather, the prosecutor must simply show that the evidence is offered to prove, e.g., the defendant's identity as the culprit under Rule 404(b)(2) or, similarly, that they had a motive, opportunity, intent, plan, etc. This would not be difficult. As a result, the question of admissibility would be left to Rule 403 and the discretion (and intuition) of the court—exactly what Rule 404 seeks to avoid.

However, notwithstanding the policies underlying Rule 404, courts regularly admit character evidence by interpreting Rule 404(b)(2) as an exception to Rule 404(b)(1).[53] Courts frequently indicate that Rule 404(b) is "one of inclusion," and that evidence offered for one of the purposes provided for in Rule 404(b)(2) is, e.g., "presumed admissible absent a contrary determination."[54] They often refer explicitly to Rule 404(b)(2) as providing for

50. FED. R. EVID. 404(b)(2).

51. *See* Daniel J. Capra & Liesa L. Richter, *Character Assassination: Amending Federal Rule of Evidence 404(b) to Protect Criminal Defendants*, 118 COLUM. L. REV. 769, 772, 776–802 (2018).

52. *See* United States v. Trenkler, 61 F.3d 45, 52 (1st Cir. 1995).

53. *See* Capra & Richter, *supra* note 51, at 778–86.

54. United States v. Williams, 796 F.3d 951, 958 (8th Cir. 2015) (quoting United States v. Wilson, 619 F.3d 787, 791 (8th Cir. 2010)).

exceptions to Rule 404(b)(1).[55] This is all to say that courts commonly admit other-acts evidence whose relevance requires character-based propensity reasoning, so long as this evidence is offered for a purpose provided for in Rule 404(b)(2).[56] A court may exclude a defendant's prior drug-trafficking offense as evidence that, because the defendant has committed a drug-trafficking offense in the past, they are more likely to have committed the offense with which they are currently charged; however, the court may well admit the evidence to prove, at least superficially, knowledge, intent, or identity—even if the evidence relies on the same reasoning.[57] Admitting the evidence for this purpose thus sanctions a character-based propensity inference to prove knowledge or intent—in violation of Rule 404(b).

For example, courts frequently admit other-acts character evidence in the form of prior incidents of harassment or discrimination to show that a defendant had a discriminatory motive or intent under Title VII and other antidiscrimination statutes.[58] In *Goldsmith v. Bagby Elevator Co.*, Goldsmith, a Black employee, sued his employer for discrimination under Title VII.[59] After the district court entered judgment on a jury verdict in favor of the employee, employer Bagby Elevator appealed.[60] Among other things, the United States Court of Appeals for the Eleventh Circuit held that under Rule 404(b), the district court appropriately admitted anecdotal "me too" evidence—evidence that Bagby Elevator (and certain common supervisors in particular) committed discrimination and retaliation against Goldsmith's coworkers.[61] Specifically, although the court held that the district court erred in admitting the evidence as "evidence of habit" under Rule 406 (the evidence did not demonstrate a sufficiently numerous and routine behavior to qualify under Rule 406), it held that the evidence "was admissible, under Rule 404(b), to prove the intent of Bagby Elevator to discriminate and retaliate."[62] The court explained:

> We have upheld the admission of coworker testimony in a sexual harassment context under Rule 404(b) to prove the defendant's "motive, . . . intent, . . . [or] plan" to discriminate against the plaintiff. Goldsmith and coworkers Jemison and Thomas were discriminated against by the same supervisor, Farley, so the experiences of Jemison and Thomas are probative of Farley's intent to discriminate. Steber was involved in the

55. *See* United States v. Sterling, 738 F.3d 228, 237 (11th Cir. 2013).

56. Capra & Richter, *supra* note 51, at 778–86.

57. *See infra* notes 66–69.

58. *See* Lisa Marshall, *The Character of Discrimination Law: The Incompatibility of Rule 404 and Employment Discrimination Suits*, 114 YALE L.J. 1063, 1065–66, 1071–74 (2005).

59. Goldsmith v. Bagby Elevator Co., 513 F.3d 1261, 1274–75 (11th Cir. 2008).

60. *Id.* at 1267.

61. *Id.* at 1285–86.

62. *Id.*

> termination decisions of all four individuals, so the experiences of Jemison, Peoples, and Thomas are probative of Steber's intent.[63]

This form of anecdotal evidence is regularly admitted in discrimination cases.[64] Some courts have acknowledged potential pitfalls associated with this evidence, but they have nevertheless held that "Rule 404(b) has come to play a significant role in employment discrimination and retaliation cases" and that "[t]he cases are basically uniform in holding as a general principle that discriminatory intent or the pretextual nature of an employment related decision may be proven by 'other acts' of discrimination or retaliation."[65]

Similarly, other-acts character evidence is frequently admitted in drug cases to prove knowledge or intent. For example, in *United States v. Manning*, the United States Court of Appeals for the First Circuit upheld the admission of evidence regarding the defendant's prior "drug dealing efforts" to prove knowledge and intent.[66] The court explained that "[t]he evidence that [the defendant] had previously sold cocaine makes it more likely both that he was aware of the contents of the plastic bags in the briefcase and that he intended to distribute the two bags of cocaine."[67] The court also highlighted that "when charges of drug trafficking are involved, this court has often upheld the admission of evidence of prior narcotics involvement to prove knowledge and intent."[68] Indeed, in the drug-trafficking context and others, many courts explicitly refer to Rule 404(b)(2) as an exception to Rule 404(b)(1) rather than as a clarification of it.[69]

---

63. *Id.* at 1286 (citations omitted). The court similarly explained the admissibility of anecdotal "me too" evidence regarding retaliation, stating that "[t]he evidence about Peoples, Thomas, and Jemison suggested that any black employee of Bagby Elevator who complained about racial discrimination was terminated." *Id.* Additionally, the court held that "Goldsmith's 'me too' evidence was also admissible, under Rule 402, as relevant to his claim of a hostile work environment." *Id.*

64. *See* Marshall, *supra* note 58, at 1065–66, 1071–74.

65. Fudali v. Napolitano, 283 F.R.D. 400, 402–03 (N.D. Ill. 2012) (citing cases).

66. United States v. Manning, 79 F.3d 212, 217 (1st Cir. 1996).

67. *Id.*

68. *Id.*; *see* United States v. Wilchcombe, 838 F.3d 1179, 1192 (11th Cir. 2016) (upholding admission of evidence that the defendant had a prior arrest for captaining a freighter containing cocaine and marijuana to prove defendant's "knowledge [in the present case] that drugs were present and that he intended to smuggle them"); *see also* United States v. Henry, 848 F.3d 1, 8–9 (1st Cir. 2017) ("Rule 404(b)(2) specifically permits the admission of a prior conviction to prove intent, and we have repeatedly upheld the admission of prior drug dealing by a defendant to prove a present intent to distribute."); 2 GEORGE E. GOLOMB & EARL JOHNSON, JR., FEDERAL TRIAL GUIDE § 23.10 (2021); Dora W. Klein, *The (Mis)Application of Rule 404(b) Heuristics*, 72 U. MIA. L. REV. 706, 722–23 (2018).

69. *See, e.g.*, United States v. Sterling, 738 F.3d 228, 237 (11th Cir. 2013) ("Rule 404(b)(1) generally prohibits the introduction of propensity evidence at trial. Rule 404(b)(2), however, provides an exception to this general rule for evidence that is also probative for some other purpose, 'such as proving motive, opportunity, intent, preparation, plan, knowledge, identity, absence of mistake, or lack of accident.'" (quoting FED. R. EVID. 404(b)(2))); *see also* United States v. Contreras, 816 F.3d 502, 510–12 (8th Cir. 2016) (involving evidence regarding prior spanking incident to prove absence of accident); United States v. Smith, 383 F.3d 700, 706–07 (8th Cir. 2004) (involving evidence of prior drug transactions to prove a drug crime); Young v. Rabideau, 821 F.2d 373, 379 (7th Cir. 1987) (involving allegations of excessive force); Klein, *supra* note 68, at 746–47.

Other-acts character evidence is also frequently admitted to prove intent, knowledge, and other mental states in fraud cases. For example, in *Turley v. State Farm Mutual Automobile Insurance Co.*, the United States Court of Appeals for the Tenth Circuit held that the district court erred in excluding evidence that the plaintiff had, in the past, conspired to defraud an insurance company.[70] The Tenth Circuit reasoned that evidence of a prior conspiracy to commit insurance fraud was relevant and admissible to prove "intent, knowledge, and absence of mistake or accident in the present proceeding."[71]

Other-acts character evidence is admitted in many other contexts also. These include, for example, arson cases, domestic abuse cases, animal abuse cases, and cases involving illegal possession of firearms and other illicit items—to name just a few.[72]

Here's the central point: In the contexts above, and in many others, courts routinely admit other-acts character evidence "to prove a person's character in order to show that on a particular occasion the person acted in accordance with the character"—in violation of Rule 404(b).[73] It is true that they admit the evidence to prove, for example, "motive, opportunity, intent, preparation, plan, knowledge, identity, absence of mistake, or lack of accident."[74] But they admit it for these purposes regardless of whether the purposes are actively disputed (other than, e.g., a general plea of "not guilty"), and they undoubtedly admit the evidence for character-based propensity reasoning.[75] They treat Rule 404(b)(2) as an exception to Rule 404(b)(1) rather than as a clarification of it.[76]

In *Goldsmith* (the Title VII case discussed above), when the court admitted evidence that Goldsmith's coworkers sustained similar harassment, discrimination, and retaliation, it was to prove that Goldsmith's employer had a *character* for doing such things—i.e., to prove that he was a racist and had a character for harassing, discriminating against, and retaliating against his Black employees—and therefore, that on the occasions in question, he acted against Goldsmith pursuant to that character. Yes, this evidence was offered to prove intent, but only through character-based propensity reasoning. It was offered

70. Turley v. State Farm Mut. Auto. Ins. Co., 944 F.2d 669, 673–74 (10th Cir. 1991).

71. *Id.* at 674.

72. *See generally* Klein, *supra* note 68, at 713–51 (discussing poor judicial reasoning surrounding Rule 404(b), including, e.g., "a fixation on fictitious 'exceptions' to Rule 404(b)").

73. FED. R. EVID. 404(b)(1).

74. FED. R. EVID. 404(b)(2).

75. *See* Capra & Richter, *supra* note 51, at 778–98 (describing a circuit split regarding the "active contest" requirement).

76. *Contra* Huddleston v. United States, 485 U.S. 681, 691–92 (1988) ("[T]he protection against . . . unfair prejudice emanates . . . from four . . . sources: first, from the requirement of Rule 404(b) that the evidence be offered for a proper purpose; second, from the relevancy requirement of Rule 402—as enforced through Rule 104(b); third, from the assessment the trial court must make under Rule 403 . . . ; and fourth, from Federal Rule of Evidence 105, which provides that the trial court shall, upon request, instruct the jury that the similar acts evidence is to be considered only for the proper purpose for which it was admitted."); *see* Capra & Richter, *supra* note 51, at 777–79.

to prove, for example, that the managers at Goldsmith's company had a track record of racism and discrimination and that, therefore, their *intent* was likely to have been discriminatory and retaliatory rather than permissible under the antidiscrimination statute.[77]

Moreover, the problem of misinterpreting Rule 404(b) is compounded by a vague doctrine that permits the admission of other-acts evidence, notwithstanding Rule 404(b), if those other acts are "inextricably intertwined" with the act at issue.[78] As the United States Court of Appeals for the Fifth Circuit has explained:

> The proper test to apply in deciding the admissibility of "similar acts" or "other acts" evidence depends upon whether the evidence in question is "intrinsic" or "extrinsic" evidence. "Other act" evidence is "intrinsic" when the evidence of the other act and the evidence of the crime charged are "inextricably intertwined" or both acts are part of a "single criminal episode" or the other acts were "necessary preliminaries" to the crime charged.[79]

Courts have arguably applied this doctrine too broadly and in a way that has caused substantial uncertainty regarding the admission of other-acts evidence:

> Although there is an obvious need for line drawing in applying Rule 404(b), many federal courts simply label uncharged offenses offered against criminal defendants as "inextricably intertwined" with the charged offense whenever they are in any way related to the charged offense. By utilizing this vague and conclusory characterization, these courts sidestep the careful Rule 404(b) analysis dictated by the Supreme Court[] . . . .[80]

Finally, courts are divided in their application of Rule 404(b). First, there is a circuit split regarding how to approach Rule 404(b) evidence. As Professors Capra and Richter explain, "The Seventh, Third, and Fourth Circuits have led a campaign to end the liberal admissibility of other-acts evidence in criminal cases by imposing limits on the prosecutorial use of such evidence."[81] Among other things, "these circuit courts have articulated a total ban on the dreaded propensity inference, barring the admission of other-acts evidence when any link in the chain of inferences supporting the relevance of the other act depends on the defendant's propensity to engage in certain conduct."[82]

---

77. Goldsmith v. Bagby Elevator Co., 513 F.3d 1261, 1286 (11th Cir. 2008).

78. Capra & Richter, *supra* note 51, at 782–83 (citing FED. R. EVID. 404(b) advisory committee's note to 1991 amendment).

79. United States v. Williams, 900 F.2d 823, 825 (5th Cir. 1990); *see* FED. R. EVID. 404(b) advisory committee's note to 1991 amendment (commenting that "[t]he amendment does not extend to evidence of acts which are 'intrinsic' to the charged offense," and citing *Williams*).

80. Capra & Richter, *supra* note 51, at 783.

81. *Id.* at 787.

82. *Id.*

Moreover, there is substantial confusion and uncertainty within circuits as well. In *United States v. Gomez*, the United States Court of Appeals for the Seventh Circuit overturned the four-part test that it (like many other circuits) applied to determine the admissibility of other-acts evidence.[83] In doing so, the court described the deteriorating state of the law under that test: "Multipart tests are commonplace in our law and can be useful, but sometimes they stray or distract from the legal principles they are designed to implement; over time misapplication of the law can creep in."[84] Reasoning that courts too frequently admit other-acts evidence without due consideration, the court held that the "four-part test for evaluating the admissibility of other-act evidence has ceased to be useful" and should be "abandon[ed] in favor of a more straightforward rules-based approach" (referring to the FRE).[85] Ultimately, the court clarified that "it's not enough for the proponent of the other-act evidence simply to point to a purpose in the 'permitted' list and assert that the other-act evidence is relevant to it"; rather, Rule 404 "allows the use of other-act evidence only when its admission is supported by some propensity-free chain of reasoning."[86] Moreover, the district court "should not just ask *whether* the proposed other-act evidence is relevant to a non-propensity purpose but *how* exactly the evidence is relevant to that purpose—or more specifically, how the evidence is relevant without relying on a propensity inference."[87]

This Seventh Circuit approach is in stark contrast to many other circuits, in which courts interpret Rule 404(b)(2) in a way that allows the rule to be arbitrarily applied and that, in many areas of the law, renders the ban on character evidence "virtually meaningless."[88]

C. Informal and Unarticulated (and Often Unrecognized) Exceptions

The problem of courts haphazardly carving out exceptions to Rule 404 is even more severe than is frequently recognized. This is because many exceptions to Rule 404 are informal and unarticulated. In many contexts, courts frequently altogether ignore the rule against character evidence. In this Subpart, I highlight a number of examples.

First, in the employment-discrimination context discussed above, courts sometimes address Rule 404(b)—generally holding that anecdotal "me too" evidence is admissible as excepted from the ban on character evidence under Rule 404(b)(2).[89] However, courts often entirely ignore the ban on character

83. *See* United States v. Gomez, 763 F.3d 845, 852 (7th Cir. 2014).
84. *Id.* at 853.
85. *Id.*
86. *Id.* at 856.
87. *Id.*
88. *Id.* at 855.
89. *See, e.g.*, Goldsmith v. Bagby Elevator Co., 513 F.3d 1261, 1286 (11th Cir. 2008).

evidence when considering the admissibility of anecdotal "me too" evidence in discrimination cases.[90]

Second, although some types of statistical evidence can constitute character evidence, courts frequently admit this evidence without identifying it as such or realizing its connection to Rule 404. Courts generally analyze the admissibility of statistical evidence under the reliability standards of *Daubert* and Rule 702 while ignoring the ban on character evidence.[91] As one author noted,

> [C]ourts encourage plaintiffs alleging discrimination to introduce statistics demonstrating the "degree of disparity between the expected and actual . . . composition of the [workforce] necessary to support an inference of discrimination," where the "composition" indicates the workers' race, sex, age, or other protected trait. Any "degree of disparity" is, however, probative of the ultimate issue in a disparate treatment case—the intention of the employer at the time she made the relevant employment decision—only insofar as it demonstrates that the employer has some enduring propensity to act in a given way.[92]

I do not mean to imply that all statistical evidence is grounded in character-based propensity reasoning. It depends on the evidence and the purpose for which it is offered. For example, statistical evidence offered to prove that a drug causes heart attacks does not rely on character-based propensity reasoning. Similarly, a descriptive study demonstrating a pay disparity between men and women is not character evidence and does not rely on propensity reasoning.

However, statistical evidence offered to prove, for example, that a pay disparity between men and women (controlling for other factors) is simply too great to be due to randomness and therefore must be due to an individual's discriminatory intent—thus supporting a claim of discrimination by a female employee against the individual—arguably may require character-based propensity reasoning. Similarly, in the antitrust context—another context in which the connection between statistical evidence and Rule 404 generally goes unnoticed in the courts (and scholarship)—statistical evidence offered to prove that pricing or output decisions of Company A and Company B (again, controlling for other factors) are simply too correlated to be due to randomness and therefore must be due to illicit collusion, thus supporting an antitrust claim by a consumer against Company A, arguably may also rely on character-based propensity reasoning.

Thus, while statistical evidence is routinely admitted in various litigation contexts, courts (and litigants) generally ignore the application of Rule 404. In most cases, it is likely that courts simply do not recognize the evidence as possible character evidence. In some cases, it is likely that the court (or litigant),

90. *See* Marshall, *supra* note 58, at 1073 (citing cases).

91. *See* FED. R. EVID. 702; Daubert v. Merrell Dow Pharms., 509 U.S. 579 (1993).

92. Marshall, *supra* note 58, at 1080–81 (quoting Berger v. Iron Workers Reinforced Rodman Local 201, 843 F.2d 1395, 1412 (D.C. Cir. 1988)).

although perhaps realizing the connection with Rule 404, simply follows precedent and applies Rule 702 without considering Rule 404. Still other courts may implicitly or explicitly apply the objective-chance doctrine, discussed *infra*, to avoid the constraints of Rule 404.[93]

Third, neither courts nor scholars have recognized the applicability of Rule 404 to sampling and representative evidence in class-action litigation. This evidence is increasingly offered as a means of avoiding individualized litigation and thereby enabling class certification. The idea is to take a sample of events, experiences, or claimants; calculate a measure of central tendency, such as a mean amount of back pay owed, a mean amount of time it takes to don and doff protective gear, or a mean damages award based on discrimination; and apply that measure to all members of a class.[94]

This method is controversial, and courts and scholars have long debated it.[95] However, the Supreme Court, lower courts, and scholars have generally ignored its relationship with Rule 404.

For example, in *Hilao v. Estate of Marcos*, the United States Court of Appeals for the Ninth Circuit upheld the decision of the district court to allow a statistical sample of claims to determine compensatory damages for a class of claims against Ferdinand Marcos, the former president of the Philippines, alleging human-rights abuses—in particular, torture, summary execution, and "disappearance"—under the former president's command.[96] A sample of 137 claims was selected randomly from a total of 9,541 claims, and a special master supervised depositions of the 137 claimants.[97] Then, based on the sample depositions, the special master not only recommended damages for the claims in the sample but also recommended damages for all of the remaining class members by extrapolating from the sample. Finally, at trial, "[t]estimony from the 137 random-sample claimants and their witnesses was introduced," and the special master testified regarding his recommendation and provided his report, which the jury had access to.[98] The jury was instructed that "it could accept, modify or reject" the recommendations in arriving at a damages award.[99]

The Ninth Circuit upheld the procedure and admissibility decisions of the lower court. However, neither the lower court nor the Ninth Circuit addressed

93. *See infra* Subpart II.D.

94. *See generally* Tyson Foods, Inc. v. Bouaphakeo, 577 U.S. 442 (2016); Hillel J. Bavli & John Kenneth Felter, *The Admissibility of Sampling Evidence to Prove Individual Damages in Class Actions*, 59 B.C. L. REV. 655 (2018); Hillel J. Bavli, *Aggregating for Accuracy: A Closer Look at Sampling and Accuracy in Class Action Litigation*, 14 LAW, PROBABILITY & RISK 67 (2015); Alexandra D. Lahav, *Bellwether Trials*, 76 GEO. WASH. L. REV. 576 (2008); Michael J. Saks & Peter David Blanck, *Justice Improved: The Unrecognized Benefits of Aggregation and Sampling in the Trial of Mass Torts*, 44 STAN. L. REV. 815 (1992).

95. *See generally* Bavli & Felter, *supra* note 94, at 659–74.

96. *See* Hilao v. Estate of Marcos, 103 F.3d 767, 782 (9th Cir. 1996).

97. *Id.*

98. *Id.* at 784.

99. *Id.*

the question of character evidence. To be sure, the applicability of Rule 404 is subtle. But, it is arguably present nevertheless. What logic permits the jury to extrapolate from the sample to determine damages for the entire class? It is through character-based propensity reasoning that the jury could be justified in extrapolating a class-wide damages award based on the sample of 137 claims. The jury must reason that the sampled claims are bound together by a common propensity of Marcos to act in a certain way, and that because this sample is representative of the entire class and the class is similarly bound together by Marcos's propensity, it is reasonable to infer damages for the class based on the sample.

This arguably constitutes character-based propensity reasoning. It is true: *logically*, perhaps this evidence does not *require* propensity reasoning. Instead, a court could, in theory, simply rely on the fact that the units—i.e., claims—are bound together statistically in some probability distribution. In theory, this fact could be sufficient for extrapolating class-wide damages. However, *legally*, this reasoning is not permissible. If it were, then presumably a court could clump together entirely independent claims not arising from the same facts or issues and award damages based on the mean damages in a small sample of those claims. This is, of course, not permissible.[100] The claimants need to be similarly situated in certain respects. They need to be bound together by something other than merely a generic probability distribution. If that something is an individual's propensity—such as in *Hilao*—then the evidence arguably relies on character-based propensity reasoning.

This idea is made clear by the reasoning in *Tyson Foods, Inc. v. Bouaphakeo*.[101] In that case, the Supreme Court held that a party could introduce representative evidence concerning the amount of time employees spent "donning and doffing protective gear" to extrapolate donning and doffing times for unsampled claims and thereby establish class-wide liability.[102] The Court highlighted that its *Tyson Foods* decision is consistent with its earlier decision in *Wal-Mart Stores, Inc. v. Dukes*, in which the Court rejected the use of sampling to establish class-wide liability.[103] The Court explained:

> The underlying question in *Wal-Mart*, as here, was whether the sample at issue could have been used to establish liability in an individual action. Since the Court held that the employees were not similarly situated, none of them could have prevailed in an individual suit by relying on depositions detailing the ways in which other employees were discriminated against by their particular store managers. By extension, if the employees had brought 1½ million individual suits, there would be little or no role for representative evidence. Permitting the use of that sample in a class action, therefore, would have violated the

100. *See generally* FED. R. CIV. P. 23(a)(3).

101. Tyson Foods, Inc. v. Bouaphakeo, 577 U.S. 442 (2016).

102. *Id.* at 446; *see also id.* at 452–60.

103. *Id.* at 457–59 (citing Wal-Mart Stores, Inc. v. Dukes, 564 U.S. 338, 367 (2011)).

> Rules Enabling Act by giving plaintiffs and defendants different rights in a class proceeding than they could have asserted in an individual action.[104]

In other words, the central issue is, what binds the claims together? Employment by a single company, Wal-Mart, does not suffice, let alone a generic probability distribution. It is particularly telling that fundamental to the Court's reasoning is the fact that the employees in *Wal-Mart* could not individually rely on "depositions detailing the ways in which other employees were discriminated against by their particular store managers."[105] In other words, because the employees were not bound together by a specific store manager, they could not establish that their claims were adequately bound together such that inferences of discrimination toward one claimant could be made from evidence of discrimination toward another. On the other hand, had they been bound together by a common manager, the evidence may have been admissible under the Court's reasoning. However, this evidence would arguably be in violation of Rule 404(b): it relies on the reasoning that a manager had a propensity to discriminate and acted in accordance with that propensity on the particular occasion at issue in a given claim. This is propensity reasoning.

Here again, not all forms of sampling evidence require propensity reasoning—far from it. Like other forms of statistical evidence, whether sampling and representative evidence constitutes character evidence depends on the particulars of the evidence and on the purpose for which it is being offered. As illustrated above, however, it can—and sometimes does—constitute character evidence.

D. Objective-Chance Evidence

In the English case *Rex v. Smith*, defendant George Smith was on trial for the murder of his wife, Bessie Mundy, whom he had recently married and who had recently inherited a substantial sum of money from her father.[106] Bessie drowned in her bathtub. The defendant alleged that he found her drowned in the bathtub, that it was an accidental drowning, and that he played no role in it.[107] To prove that the defendant killed Bessie and that she did not die of an *accidental* drowning, the prosecutor sought to introduce evidence that two subsequent wives of the defendant also drowned in their bathtubs. The court admitted the evidence.[108] The court instructed the jury that it may not use the evidence for the inference that the defendant is of murderous character and therefore likely to have committed murder. Rather, they were to use it as

104. *Id.* at 458.

105. *Id.*; *see also* Hillel J. Bavli, *Sampling and Reliability in Class Action Litigation*, 2016 CARDOZO L. REV. DE NOVO 207, 214–19.

106. Rex v. Smith (1915) 84 LJKB 2153; *see* FISHER, *supra* note 19, at 195–98.

107. *See Smith*, 84 LJKB at 2154.

108. *Id.*

evidence of whether Bessie's death was accidental or by the design of the defendant.[109] The defendant was found guilty of murdering Bessie.[110]

This famous case is commonly cited as the seminal case for the doctrine of objective chances.[111] The doctrine, often viewed as involving a non-propensity use of other-acts evidence, allows the introduction of other occurrences similar to an event at issue to show that there is a low probability that the event was an accident.[112] For example, in *Smith*, the court admitted evidence of the purportedly accidental deaths of the defendant's subsequent wives to show not directly that the defendant had a murderous character and acted in line with that character but that it was too improbable for three of the defendant's wives to drown accidentally.

The doctrine of chances is not limited to literal showings of absence of mistake or accident. Rather, it has been applied to prove various elements in various types of claims—including to show discriminatory motive or intent in discrimination cases, knowledge or intent in drug-trafficking cases, absence of accident in arson cases, absence of accident in child abuse cases, and others.[113]

This doctrine, although frequently characterized as involving a non-propensity use of other-acts evidence, is the subject of substantial controversy and debate. Proponents of the doctrine, such as Professor Imwinkelried, argue that evidence falling within the doctrine of objective chances does not rely on propensity reasoning because it does not require the jury to "consciously advert to the question of the accused's personal, subjective bad character."[114] According to Professor Imwinkelried, impermissible reasoning under Rule 404(b) entails the following chain of inferences: "[t]he accused's other misdeed(s)" >> "[t]he accused's subjective bad character" >> the conclusion that "[o]n the occasion alleged in the pleadings the accused acted 'in character,' consistently with his or her subjective bad character."[115] According to Professor Imwinkelried, the reasoning underlying the doctrine of chances, however, involves a different chain of inferences: "[o]ther incidents the accused was involved in" >> "[t]he objective improbability of so many accidents (an extraordinary coincidence)" >> the conclusion that "[o]ne or some of the incidents were not accidents."[116]

---

109. *Id.* at 2154, 2155–56; *see* FISHER, *supra* note 19, at 195–98; *see also* United States v. Woods, 484 F.2d 127, 134–35 (4th Cir. 1973) (holding, in case involving the death of an infant after episodes of cyanosis, that evidence that numerous other children in the defendant's care suffered from cyanotic episodes, some of whom died following these episodes, was admissible to prove that the victim's death was by design rather than due to accident or natural causes).

110. *Smith*, 84 LJKB at 2153; *see also id.* at 2157.

111. Imwinkelried, *supra* note 8, at 6–8.

112. *Id.* at 2–3, 6–7; FISHER, *supra* note 19, at 195–201.

113. *See* Imwinkelried, *An Evidentiary Paradox*, *supra* note 25, at 419–25.

114. Imwinkelried, *supra* note 8, at 7; *see* JOHN HENRY WIGMORE, A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW § 302 (2d ed. 1923).

115. Imwinkelried, *supra* note 8, at 5.

116. *Id.* at 7.

According to this argument, the policy concerns invoked by the ban on character evidence are not invoked when the "intermediate inference is the objective improbability of so many accidents"—reasoning that, according to this argument, does not rely on inferences regarding the "accused's personal, subjective bad character."[117]

Moreover, this argument posits that while "[u]nder a character theory, the jurors must use the accused's personal, subjective character as a predictor of conduct on a particular occasion," objective-chance evidence requires only that jurors "use their common sense to determine which contention is more plausible—the defense's contention that all the incidents are accidents or the prosecution's contention that at least one or some of the incidents amount to crimes."[118] Professor Imwinkelried argues that because these inferential steps do not involve resort to impermissible propensity reasoning—that a party has a certain type of bad character and is likely to have acted in accordance with that character on a particular occasion—objective-chance evidence does not constitute character evidence.[119]

On the other hand, opponents of the doctrine, such as Professor Rothstein, have argued (among other criticisms of the doctrine) that, notwithstanding the disguised form of objective-chance evidence, reliance on character-based propensity reasoning is "inescapable."[120] They reason as follows:

> The essence of this probable guilt argument [i.e., the argument above suggesting that objective-chance evidence does not require propensity reasoning] is that there is a disparity between the chances, or probability, that an *innocent* person would be charged so many times and the chances, or probability, that a *guilty* person would be charged so many times. If there is such a disparity, however, it is only because a guilty person would have the *propensity* to repeat the crime. If it were not for the propensity to repeat, the chances, or the probability, that an innocent person and a guilty person would be charged repeatedly would be identical. Hence, the argument hinges on propensity and runs afoul of the first sentence of Rule 404(b). The effort to reconcile the permission in the Rule with the prohibition in the Rule has failed.[121]

Opponents of the doctrine of chances have also argued that even if objective-chance evidence does not require propensity reasoning, the reasoning required by the evidence would be too similar to propensity evidence and therefore should be excluded under Rule 403 since we could not expect a jury

117. *Id.*

118. *Id.*

119. *See id.*

120. Rothstein, *Intellectual Coherence in an Evidence Code*, *supra* note 25, at 1261.

121. *Id.* at 1262–63; *see also* Marshall, *supra* note 58, at 1081–82; Andrew J. Morris, *Federal Rule of Evidence 404(b): The Fictitious Ban on Character Reasoning from Other Crime Evidence*, 17 REV. LITIG. 181, 200 n.74 (1998).

to apply permissible reasoning without also applying impermissible propensity reasoning.[122]

Professor Imwinkelried disagrees with Professor Rothstein and others who argue that the doctrine of chances necessarily leads to propensity reasoning. He reasons that "[t]he doctrine's applicability does not dictate the inference that all of the incidents are non-accidental or even that the charged incident is non-accidental."[123] Rather, he argues, "The only warranted inference from the doctrine's applicability is that one or some of the incidents are likely not accidents."[124] Thus:

> On the one hand, the doctrine of chances evidence is logically relevant and presumptively admissible; to a degree, the negative disproof of the random chance hypothesis affirmatively increases the probability of the competing explanations for the outcomes, including the explanations which do not entail propensity inferences. On the other hand, the doctrine of chances standing alone might be legally insufficient to sustain the prosecution's or plaintiff's burden of production.[125]

The debate continues. In the meantime, however, the doctrine of chances is frequently misused to justify the admission of character evidence and thereby inadvertently propagate uncertainty and incorrect law. This occurs in various ways, including overbroad applications of the rule in cases in which ordinary character evidence is superficially masked as evidence offered to prove absence of accident;[126] misinterpretations of Rule 404(b)(2) as providing for exceptions to the ban on character evidence for evidence offered to prove "absence of mistake" or "lack of accident;"[127] and poorly reasoned carve-outs from Rule 404(b)(1).[128] Moreover, the doctrine, as currently defined, is frequently ignored or avoided in circumstances in which the doctrine would apply were it defined explicitly as an *exception* to the rule against character evidence.[129]

II. AN AGGREGATION-EVIDENCE MODEL OF OTHER-ACTS CHARACTER EVIDENCE

In this Part, I apply scientific principles of estimation to model the effect of character evidence on accuracy. Specifically, I build on my previous work to

122. *See* Morris, *supra* note 121, at 200 n.74; *see also* Imwinkelried, *supra* note 8, at 13.

123. Imwinkelried, *supra* note 8, at 10.

124. *Id.*; *see also* Imwinkelried, *An Evidentiary Paradox*, *supra* note 25, at 456.

125. Imwinkelried, *An Evidentiary Paradox*, *supra* note 25, at 456.

126. *See* People v. Burnett, 2 Cal. Rptr. 3d 120, 130–32 (Ct. App. 2003) (examining the admissibility of evidence that defendant "beat a stray dog to death" to rebut defendant's accident defense); FISHER, *supra* note 19, at 194; *see also supra* notes 68–69.

127. *See supra* Subpart I.B.

128. *See supra* Subpart I.B.

129. *See infra* Part III.

examine why courts depart from the rule against character evidence and to understand the role of objective-chance evidence in these departures. I apply the aggregation-evidence model to define objective-chance evidence and analyze the unique nature of this type of evidence in terms of its ability to improve accuracy.

A. A Model of Accuracy

Assume an individual employee brought an action against his employer based on the employer's failure to pay overtime wages, as required by statute, for time that the employee spent donning and doffing protective gear.[130] Assume that there is a correct amount of compensatory damages that could, in theory, be computed formulaically based on the precise amount of overtime hours for which the plaintiff should have been paid. However, in practice, we do not know the precise amount of time that the employee spent donning and doffing protective gear; therefore, we do not know the precise amount of compensatory damages to award. We need to make inferences from incomplete facts and estimate it.

More generally, let us define a correct judgment regarding a fact or outcome (such as a damages award, liability finding, or sentence) as the judgment that would result from perfect information regarding a case—including all facts, norms, and law related to it.[131] However, we do not have perfect information regarding a case; therefore, we need to infer missing facts and estimate the correct outcome.[132] Let us then define *error* in terms of the distance between the estimate and the correct judgment and *accuracy* in terms of their proximity.[133]

Using standard definitions of proximity and distance in statistics, we can deconstruct error into two components: *variance* and *bias*. Variance is a measure of dispersion. As dispersion around a mean, or "expected," value increases, variance increases, and vice versa.[134] "Precision" is the inverse of variance: as dispersion increases, precision decreases, and vice versa.[135] Bias, on the other hand, measures the difference between the correct judgment and the expected judgement, or "expectation"—that is, the mean of repeated samples of the judgment, or repeated adjudications.[136] If the estimator is, on average, equal to

130. *See generally* Tyson Foods, Inc. v. Bouaphakeo, 577 U.S. 442 (2016).

131. Hillel J. Bavli, *The Logic of Comparable-Case Guidance in the Determination of Awards for Pain and Suffering and Punitive Damages*, 85 U. CIN. L. REV. 1, 12–13 (2017) [hereinafter *The Logic of Comparable-Case Guidance*]; Bavli, *supra* note 94, at 74–78. We can similarly define a *distribution* of correct judgments that reflects various sources of inherent uncertainty (e.g., in the law); for simplicity, however, let us assume a single correct judgment. *Id.* at 13 n.50.

132. *Id.* at 13.

133. *Id.*

134. *See id.* at 14–15.

135. *See id.* at 14.

136. *Id.*

the correct outcome—that is, if the mean of repeated samples of the estimation is equal to the correct outcome—then it is "unbiased." If not, then it is "biased."[137]

It is preferable for a judgment to be unbiased. But accuracy is based on bias and variance, and even if a judgment is correct on average, it may be highly variable around the correct judgment and thereby involve a high degree of error.[138] For example, assume that in a certain case, the correct value of a damages award is $100,000. Then:

> [R]epetitions of an unbiased adjudication may generate estimate values (i.e., damage awards) of $0, $50,000, $150,000, and $200,000, which are indeed centered at the correct value of $100,000; however, the awards are highly dispersed around $100,000. We would, for example, prefer that repeated adjudications generate the values $90,000, $95,000, $105,000, and $110,000; or even better, $100,000, $100,000, $100,000, and $100,000.[139]

Frequently, we are interested in minimizing error and maximizing accuracy—not in minimizing bias or variance in particular. Therefore, it is important to consider bias and variance together. Indeed, there is often a *trade-off* between bias and variance. This means that in order to reduce variance, it is frequently necessary to introduce some bias. In statistics, this is known as a *bias–variance trade-off*.[140] For example, in a recent Article, I have argued that in order to address the vast unpredictability of awards for pain and suffering and punitive damages, courts should consider providing jurors with information regarding awards in factually similar cases.[141] Although this method may add a small amount of bias, it would greatly reduce the variability of the awards and thereby generate very substantial accuracy benefits.[142]

For example, assume that the correct damages award in a case is $100,000. We may well prefer a trial process that, upon repetition, would produce awards of $90,000, $93,000, $97,000, and $100,000 over a trial process that, upon repetition, would produce awards of $0, $50,000, $150,000, and $200,000.[143] The second trial process is unbiased but involves a high degree of variability, whereas the first trial process is biased but involves only a small degree of bias and far less variability.[144] Even the most incorrect award from the first, biased process ($90,000) entails greater accuracy than the least incorrect awards from the second, unbiased process ($50,000 and $150,000).

---

137. *Id.*

138. *See id.* at 14–15.

139. *Id.* at 14.

140. *See An Aggregation Theory of Character Evidence*, *supra* note 17, at 46.

141. *See The Logic of Comparable-Case Guidance*, *supra* note 131, at 15–24.

142. *See* Hillel J. Bavli & Reagan Mozer, *The Effects of Comparable-Case Guidance on Awards for Pain and Suffering and Punitive Damages: Evidence from a Randomized Controlled Trial*, 37 YALE L. & POL'Y REV. 405, 406–09, 412–19, 430–50 (2019); *The Logic of Comparable-Case Guidance*, *supra* note 131, at 15–24.

143. Bavli & Mozer, *supra* note 142, at 416.

144. *Id.*

Thus, assume that α is the correct award, sentence, factual determination, or other judgment to be decided in a case. Assume that $\hat{\alpha}$, a random variable, is the estimate of the correct judgment—the *actual* judgment. α is the "estimand"—the thing we want to estimate—whereas $\hat{\alpha}$ is the "estimator"—the thing we use to estimate it. Let $E(\hat{\alpha})$ represent the expectation of $\hat{\alpha}$. $\hat{\alpha}$ will therefore equal $E(\hat{\alpha})$ on average. Then, $Bias = E(\hat{\alpha}) - \alpha$. We say that $\hat{\alpha}$ is unbiased if $E(\hat{\alpha}) = \alpha$.[145] Further, let $V(\hat{\alpha})$ represent the variance of $\hat{\alpha}$, where $V(\hat{\alpha}) = E\big(\hat{\alpha} - E(\hat{\alpha})\big)^2$. Finally, let us define the error associated with $\hat{\alpha}$ using the standard statistical measure of *mean squared error* (*MSE*), where $MSE(\hat{\alpha}) = E[(\hat{\alpha} - \alpha)^2]$. In other words, the error associated with an actual judgment is equal to the expected square difference between the actual judgment and the correct value.[146] Based on these definitions, it can be confirmed that the error associated with a judgment can be deconstructed into variance and bias. That is, $MSE(\hat{\alpha}) = E[(\hat{\alpha} - \alpha)^2] = Variance + Bias^2$.[147]

B. Judgment Variability and Event Variability

Let us now build on the model of accuracy described in the previous Subpart to better understand the role of information aggregation—that is, the role of using information regarding other events (defined formally in the following Subpart)—in generating accurate judgments. Assume that instead of an individual action alleging damages based on unpaid overtime wages, a class of employees sued their employer for overtime wages in the form of a class or collective action, as in *Tyson Foods, Inc. v. Bouaphakeo*.[148] In *Tyson Foods*, as in the example above, the claimant employees lacked records of the precise donning and doffing times for which they were owed overtime wages.[149] This stood as an impediment to class certification under Federal Rule of Civil Procedure 23(b)(3), which requires that common issues predominate over individual issues.[150] The Supreme Court, however, approved of the use of sampling evidence to prove the amount of time that each member of the putative class took to don and doff protective gear.[151] To show that it would be unnecessary to individually litigate the time that it took for each member of the class to don and doff protective gear, the putative class sought to introduce evidence of a sample of donning and doffing times and, particularly, a sample mean (e.g., of 18 minutes per day to don and doff protective gear) to allow the jury to

145. *The Logic of Comparable-Case Guidance*, *supra* note 131, at 14.

146. *Id.* at 15. MSE is a convenient and well-accepted definition of error for various reasons. Note, however, that my argument holds under other accepted definitions.

147. *Id.*

148. Tyson Foods, Inc. v. Bouaphakeo, 577 U.S. 442, 447–48 (2016).

149. *Id.* at 450.

150. FED. R. CIV. P. 23(b)(3).

151. *Tyson Foods*, 577 U.S. at 452–60.

extrapolate from that sample (and sample mean) for each member of the putative class.[152]

In light of multiple claims (or sets of facts) of the same general kind, such as those in *Tyson Foods*, estimates associated with legal claims can be further characterized using a "hierarchical" model. The damages outcome associated with each claim involves two levels of variability. The first level is *claim variability*, defined as the variability associated with the *correct* claim outcomes (e.g., the variability in the correct awards—reflecting true donning and doffing times—from claim to claim in *Tyson Foods*). The second level is *judgment variability*, defined as the variability (or uncertainty) associated with the *estimates* of each claim given the respective claim's correct outcome (e.g., the variability associated with the findings of fact regarding each individual claimant's donning and doffing times given a claimant's true donning and doffing times).[153]

Note that these concepts easily extend beyond donning and doffing times to inferences surrounding information aggregation generally. Inferring information (a fact or outcome) associated with a particular event at issue in a case from information associated with distinct events of the same general kind involves two types of variability: *event variability* and *judgment variability*. Event variability refers to variability in the actual features of the events, whereas judgment variability refers to variability arising from the estimation (or inference) regarding the true characteristics of the events.[154]

More formally, assume that there are n claims in a class and that we index all of the claims using $i = 1, 2, 3 \ldots n$. Assume that $\hat{\alpha}_i$, the actual adjudicated outcome (or event judgment generally) associated with any particular claim i in the class, is unbiased but variable around α_i, the correct outcome (or event) associated with that claim.[155] In particular, $\hat{\alpha}_i$ is "distributed" with mean α_i and variance σ^2, where σ^2 is the judgment variability associated with claim i. Notationally, $\hat{\alpha}_i \sim (\alpha_i, \sigma^2)$.[156]

Judgment variability arises from randomness associated with a judgment—for example, randomness in the selection of jurors (where one jury may award $0 while another awards $10,000 and another awards $100,000), the selection of a judge, the selection of attorneys, and the details associated with the presentation of evidence. If a claim were adjudicated repeatedly and independently under different conditions (e.g., different juries, judges,

152. *Id.* at 450, 454. The expert for the putative class "averaged the time taken in the observations to produce an estimate of 18 minutes a day for the cut and retrim departments and 21.25 minutes for the kill department." *Id.* at 450.

153. *The Logic of Comparable-Case Guidance*, *supra* note 131, at 18; Bavli, *supra* note 94, at 75–78, 81–83. *See generally* Edward K. Cheng, *When 10 Trials Are Better than 1000: An Evidentiary Perspective on Trial Sampling*, 160 U. PA. L. REV. 955 (2012); Saks & Blanck, *supra* note 94.

154. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 47.

155. Note, it is not necessary to assume unbiasedness; however, it is convenient for explanatory purposes.

156. *The Logic of Comparable-Case Guidance*, *supra* note 131, at 18.

attorneys, etc.), judgment variability would be equal to the variance of all of the outcomes associated with the repeated judgments.[157] This concept applies to factual determinations, verdicts, damage awards, sentences, and other judgments. Importantly, judgment variability is directly related to the evidence in a case. Stronger evidence leads to less judgment variability and vice versa.

Event variability is different. It represents differences in the true facts or correct outcomes associated with a group of events or claims. For example, in *Tyson Foods*, each claim involves a different donning and doffing time. Still assuming that the correct damages award for each claim is a function of the claim's true donning and doffing time, and assuming that donning and doffing time is the only factor that distinguishes one claim from another, claim variability—a particular type of event variability—is equal to the variance of the correct damage awards, which reflects the true donning and doffing times for the claims.[158] Let us denote event variability by τ^2. We can then write $\alpha_i \sim (\mu, \tau^2)$, meaning that the correct damage awards, α_i, are distributed around some central mean, μ, with variability (measured in terms of variance) τ^2.[159] Note again that event variability can refer to variability in events other than claims. For example, it can refer to the variability associated with uncharged prior assaults or incidents of discrimination that have not materialized into legal claims. I will utilize the τ^2 notation to denote both event variability and claim variability since the former is just a generalization of the latter.

C. The Impact of Character Evidence on Accuracy

In a recent Article, I describe *aggregation evidence* as a type of evidence that "involves inferring information about an event at issue in a case from information about distinct events of the same general kind."[160] This category of evidence includes character evidence as well as various other types of evidence. For example, it includes various forms of anecdotal evidence (e.g., offered to prove discriminatory intent in a Title VII case), sampling evidence in employment cases, various types of statistical evidence, and "comparables" evidence used to review a criminal sentence or damages award or to prove the value of property confiscated by the government in a takings case.[161]

For example, other-acts character evidence involves inferring information regarding an event at issue in a case—e.g., whether the defendant committed a robbery—from information regarding distinct events of the same general

157. *See* Bavli, *supra* note 94, at 75–78; Saks & Blanck, *supra* note 94, at 833–37. For simplicity, I assume that judgment variability is equal for all claims in a class.

158. *See The Logic of Comparable-Case Guidance*, *supra* note 131, at 18.

159. *Id.*; Bavli, *supra* note 94, at 81–83.

160. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 45.

161. *Id.*

kind—e.g., information regarding prior robberies that the defendant is alleged to have committed.[162]

The various forms of aggregation evidence are bound together by the inferences that they involve and by the mechanisms through which they affect accuracy.[163] Specifically, whether a particular piece of aggregation evidence improves the accuracy of a judgment depends on the bias–variance trade-off that it involves. For example, providing a jury with information regarding a robbery defendant's prior alleged robberies biases the jury's judgment regarding the defendant's role in the robbery at issue. However, because the other-acts evidence is probative—it affects, via propensity reasoning, the likelihood that the defendant committed the robbery in question—it reduces the unpredictability of the judgment.

In line with the model above, if our goal is to improve the accuracy of a judgment $\hat{\alpha}_*$, where $\hat{\alpha}_* \sim (\alpha_*, \sigma^2)$ and $\alpha_* \sim (\mu, \tau^2)$, then aggregation evidence can be understood as other samples, $\alpha_i \sim (\mu, \tau^2)$, for $i = 1. \ . \ . n - 1$, from the same distribution from which α_* arises, offered to inform $\hat{\alpha}_*$ (i.e., to provide information regarding α_*) through information about μ and τ^2.[164] For example, if α represents whether a defendant in a discrimination case had discriminatory intent, with $\alpha_* = 1$ indicating that there was discriminatory intent and $\alpha_* = 0$ indicating that there was not, then the distribution from which α arises is characterized by the parameter $\mu = p$, which indicates the probability that $\alpha_* = 1$.[165] If $p = 0.8$, for example, then there is a high probability of discriminatory intent.[166] In this scenario, other-acts evidence—e.g., anecdotal evidence regarding prior accusations of discrimination—provides information regarding α, whether or not there was discriminatory intent in the current case, via information regarding p, the parameter that characterizes the distribution from which α arises.[167]

Now, whether aggregation evidence improves accuracy—and specifically, whether its accuracy-enhancing, downward effect on judgment variability dominates its accuracy-reducing, upward effect on bias—is based on two factors. First, it depends on the precision of the aggregation evidence in terms of both the factual uniformity of the evidence (relative to itself and relative to the act at issue) and the strength of the evidence.[168] For example, weak evidence of one instance of alleged prior discrimination provides far less precision-enhancing benefit than strong evidence regarding multiple instances of alleged prior discrimination. Further, the more uniform the prior instances—with

162. *Id.*
163. *Id.*
164. *Id.* at 50.
165. *Id.* at 51.
166. *Id.*
167. *Id.*
168. *Id.*

respect to each other and with respect to the incident at hand—the more useful the evidence will be.[169]

Second, it depends on the unpredictability of the judgment without the aggregation evidence.[170] For example, aggregation evidence in the form of other-award information would be far more useful in guiding a pain-and-suffering award than a damages award for medical expenses because there is likely little evidence and much unpredictability surrounding the former determination while there is likely abundant evidence and little unpredictability surrounding the latter determination.[171]

Again, existing unpredictability is a function of the evidence in a case in the absence of the aggregation evidence. If there are clear records of medical expenses, this determination will be relatively predictable, and aggregation evidence regarding medical expenses in other cases will not improve accuracy—it will introduce bias while not causing a substantial reduction in variability. On the other hand, if there is little evidence to guide a jury in determining a pain-and-suffering award, aggregation evidence can fill the gap and provide necessary guidance—that is, the beneficial effect with respect to unpredictability is likely to outweigh any introduction of bias.[172]

Importantly, the concepts of bias and variance and the impact of aggregation evidence on accuracy can be described in terms of probative value and prejudicial effect under Rule 403. In *An Aggregation Theory of Character Evidence*, I argue that the general ban on character evidence can be justified by the model above, given the effects of other-acts character evidence on bias and precision. I argue that this is especially so in light of a concept that I refer to as "policy-disfavored bias"—that is, "bias to which there is a particular sensitivity or aversion beyond simply its implications for accuracy."[173] For example, we would be particularly reluctant to allow bias based on race or gender. Similarly, we would be reluctant to allow bias that disadvantages a defendant in a criminal case—even if such bias gives rise to accuracy benefits.[174] I argue that character-based bias can be understood as a form of policy-disfavored bias. We prefer to give people the benefit of the doubt and assume that, regardless of their past conduct, they may have improved generally or otherwise acted differently in the instance in question. As many have highlighted previously, other-acts character evidence may well be probative, but we have a particular aversion to bias arising from character-based generalizations.[175]

169. *Id.* at 51–52.
170. *Id.* at 52–53.
171. *Id.* at 49.
172. *See generally The Logic of Comparable-Case Guidance*, *supra* note 131.
173. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 53.
174. *Id.*
175. *Id.* at 54.

I argue that the law's aversion to character-based bias "translates to an uneven weighting of bias and variance with respect to accuracy" and that this explains the decision to replace a case-by-case application of a Rule 403 balancing with Rule 404's assumption that the probative value of character evidence will be substantially outweighed by the risk of unfair prejudice.[176] At the same time, however, it arguably also provides an explanation for Rule 404's many exceptions: "the rule against character evidence has replaced a balancing that reflects a trade-off between variance and bias—even heavily weighted policy-disfavored bias—with a ban that altogether excludes character evidence regardless of the trade-off that exists in a particular case."[177]

In that Article, I therefore ask whether there are types of other-acts character evidence for which the bias–variance trade-off is so favorable that the evidence categorically improves accuracy, even accounting for the law's particular aversion to character-based bias. I argue that the answer is yes—this occurs for objective-chance evidence.[178]

D. Defining Objective-Chance Evidence

Define objective-chance evidence as "evidence regarding events of the same general kind as the event at issue, offered to prove that the event at issue, in light of the number of similar events evidenced, did not occur randomly but rather occurred in accordance with the events evidenced."[179] Applying the model above, let the truth regarding the event at issue be represented by α_* and the truth regarding the other events evidenced be represented by α_i. Let $\alpha_* \sim (\mu, \tau^2)$ and $\alpha_i \sim (\mu, \tau^2)$, and assume that each takes values 0 or 1. Also, assume that the party offering the other-acts evidence aims to prove $\alpha_* = 1$. For example, the prosecutor in *Rex v. Smith* offered other-acts evidence to prove that the defendant murdered Bessie, represented by $\alpha_* = 1$. Objective-chance evidence is evidence of other events α_i, for $i = 1 \ldots n-1$, all arising from the same distribution as α_* (i.e., $\alpha_i \sim (\mu, \tau^2)$), offered by the sponsoring party to show that $\alpha_* = 0, \alpha_1 = 0, \alpha_2 = 0, \ldots$, and $\alpha_{n-1} = 0$ would be very improbable, given μ, the probability of $\alpha_i = 0$.[180] In *Smith*, the prosecutor offered evidence that two of the defendant's subsequent wives similarly died of drowning to prove that it is very improbable that three of the defendant's wives would drown in separate incidents by accident, without any of them being murdered by the defendant—i.e., to prove that $\alpha_* = 0, \alpha_1 = 0,$ and $\alpha_2 = 0$ is

176. *Id.*
177. *Id.*
178. *Id.*
179. *Id.* at 54–55.
180. *Id.* at 55; *see* Rex v. Smith (1915) 84 LJKB 2153, 2154.

highly improbable—in light of the very low probability μ of someone drowning accidentally in the bathtub.[181]

This reasoning is distinct from that associated with ordinary character evidence. Ordinary character evidence can be understood in statistical terms as providing information regarding a "prior" probability distribution.[182] For example, evidence regarding a defendant's prior robberies, offered to prove that the defendant has a propensity to commit robberies and is therefore more likely to have committed the robbery in question, can be understood as replacing an initially "noninformative" prior for committing robberies—one that provides no particular information regarding the defendant's propensity for committing robbery—with an understanding of the defendant's propensity that incorporates information regarding the defendant's prior acts. This updated understanding of μ increases the likelihood that the defendant committed the robbery in question—it increases the probability that $\alpha_* = 1$.[183]

The chain of inferences associated with objective-chance evidence is different. In *Smith*, the prosecutor sought to introduce evidence of the subsequent drownings to show as follows: it would be extremely unlikely that all three drownings were due to accident; therefore, at least one of the drownings was likely by the defendant's design; therefore, it is more likely that the defendant committed the murder in the case at hand.[184] At the center of this chain of inferences is a hypothesis test: it is so unlikely that all three drownings occurred by chance that the jury should replace the null hypothesis of accidental drownings, or randomness, with an alternative hypothesis that reflects the defendant's propensity for murder.[185]

In terms of prior probabilities, we similarly start with a "noninformative" prior that provides no information regarding the defendant's propensity for murder.[186] But then, the information regarding the previous drownings is not simply incorporated into the prior. Rather, contrary to ordinary propensity reasoning, which would view each prior incident as relevant to propensity and the likelihood that the defendant committed the crime in question, objective-chance evidence involves a hypothesis test.[187] This chain of inferences assumes that each prior drowning is not necessarily relevant. Rather, each event is relevant only in combination with the other drowning events and only if the combination of the similar events is sufficiently unlikely to occur by chance that

181. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 55.

182. *Id.* at 55–57. I use Bayesian reasoning to define and characterize character evidence, and objective-chance evidence in particular.

183. *Id.* at 55–56.

184. *Id.* at 56–57; *see Smith*, 84 LJKB at 2153–54.

185. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 56–57.

186. *Id.* at 55–56.

187. *Id.*

it justifies rejecting the hypothesis of accidental drownings.[188] If and only if this is the case, then the understanding of μ is updated with information regarding the previous drownings and the defendant's propensity to commit murder. In turn, the updated understanding of the prior suggests an increased probability that the defendant murdered Bessie (i.e., that $\alpha_* = 1$).[189]

E. The Unique Nature of Objective-Chance Evidence

As discussed above, aggregation evidence has an especially strong potential for improving the accuracy of a judgment when judgment variability is high in the absence of the aggregation evidence and the precision of the aggregation evidence itself is high.[190] This makes sense: aggregation evidence is most useful when there is a lack of other evidence, resulting in high levels of unpredictability, and the aggregation evidence itself provides a strong informational signal and thereby has a high probability of effectively filling the evidentiary gap. Objective-chance evidence has both of these features.[191]

First, objective-chance evidence informs matters for which there is generally little evidence. Judgment variability—a function of the evidence in a case—is therefore high. Specifically, as defined herein, objective-chance evidence is usually offered to prove a party's state of mind. For example, it is frequently offered to prove intent, motive, or knowledge.[192] However, proving a party's state of mind is extremely difficult. For example, there is little evidence at a plaintiff's disposal to prove the discriminatory intent of an employer. It requires understanding the employer's motivations—their mental operations. It is frequently proved circumstantially using anecdotal evidence or statistical evidence.[193] Similarly, there is little evidence at a prosecutor's disposal to prove a defendant's knowledge or intent in a drug-trafficking case. Again, this involves proving a defendant's state of mind. It is often similarly accomplished using prior acts.[194]

Furthermore, even in a case like *Smith*, which involves a question of the defendant's actions rather than mental state, objective-chance evidence is generally offered not to prove *who* committed a crime but rather *whether* there

188. *Id.* at 56–57
189. *Id.*
190. *Id.* at 57.
191. *Id.*
192. *Id.*
193. *Id.*
194. *Id.*

was even a crime in the first place.[195] This question is again extremely difficult to evidence.[196]

In summary, objective-chance evidence is generally offered to prove matters for which there is little other available evidence. This results in high levels of unpredictability in the absence of the objective-chance evidence—that is, high levels of judgment variability.

Second, objective-chance evidence involves particularly high levels of precision relative to other forms of character evidence.[197] The hypothesis test at the center of the chain of inferences associated with objective-chance evidence ensures this: rejecting the null hypothesis of randomness requires a combination of sample size (i.e., that there be sufficiently numerous incidents), internal uniformity (i.e., that the prior events be similar to one another), uniformity relative to the event at issue, and low levels of uncertainty surrounding the occurrence of the prior events.[198] All of these factors ensure a high level of precision to constitute objective-chance evidence and logically enable the rejection of a null hypothesis of randomness or chance.[199]

For example, to prove that a defendant started a fire purposefully in an arson case, a prosecutor may offer evidence of destructive fires in the defendant's previous homes. The evidence involves multiple discrete, binary, low-probability events involving relative certainty regarding whether each event occurred.[200] Moreover, the events are uniform—and extremely uniform with respect to the informational signal that they entail for objective-chance evidence—with respect to one another and with respect to the event at issue.[201]

Thus, the combination of "informing matters that entail high judgment variability and providing precise information about those matters" allows objective-chance evidence to stand apart from other types of evidence—and other types of character evidence in particular—in its ability to improve accuracy.[202] Consequently, "it can often seem counterintuitive, or in certain circumstances even absurd, to exclude this evidence."[203] This has resulted in

195. *See* Rex v. Smith, 1915 LJKB 2153; *see also* United States v. Woods, 484 F.2d 127, 134–35 (4th Cir. 1973) (affirming the admission of evidence regarding the cyanotic episodes of other children in the defendant's care to prove that the death of an infant in her custody was by her design rather than due to accident or natural causes).

196. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 57.

197. *Id.* at 57–58.

198. *Id.*

199. *Id.*

200. *Id.*

201. *Id.*

202. *Id.* In a recently published article, Professor Steven Goode has asserted a similar argument: "Where intent is truly a controverted issue, Rule 404's categorical judgment that the risk of unfair prejudice outweighs the probative value of other-acts evidence—even when it requires a character-propensity inference—is probably wrong." Steven Goode, *It's Time to Put Character Back into the Character-Evidence Rule*, 104 MARQ. L. REV. 709, 802 (2021). He thus concludes that Rule 404(b) should include a "true exception" for other-acts character evidence that is offered to prove intent "where intent is controverted." *Id.* at 807.

203. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 57–58.

substantial pressure on courts to carve out exceptions to the rule against character evidence or even to ignore it altogether.[204]

III. AN OBJECTIVE-CHANCE EXCEPTION TO THE BAN ON CHARACTER EVIDENCE

As summarized in the previous Part, in *An Aggregation Theory of Character Evidence*, I introduced a model of character evidence as a form of aggregation evidence, which ties together evidence from various contexts based on the inferences and mechanisms through which it affects accuracy. I then applied the concept of a bias–variance trade-off to discuss both the impact of character evidence on accuracy and the unique nature of objective-chance evidence. In this Part, I apply my analysis to develop a more logical approach to objective-chance evidence and a more coherent rule against character evidence. Specifically, I begin this Part by applying the aggregation-evidence model to address longstanding debate and confusion surrounding the nature of objective-chance evidence. I argue that although objective-chance evidence can technically be understood as a non-propensity use of other-acts evidence, it is *best* understood as a form of character evidence, requiring propensity reasoning. Then, based on my analysis, I argue that an objective-chance exception to the rule against character evidence may allow a more logical and effective approach to objective-chance evidence and character evidence more broadly.

A. Objective-Chance Evidence as a Particular Form of Character Evidence

As discussed in Part I, there is longstanding disagreement regarding whether objective-chance evidence depends on character-based propensity reasoning. The model and analysis above point to the conclusion that it does. More specifically, however, although propensity reasoning is not absolutely *necessary* for the relevance of the evidence, it is necessary to achieve the primary and overwhelming probative value of the evidence. The entire purpose of the evidence is to reject a hypothesis of chance or accident in favor of an understanding based on the defendant's particular propensity.

Arguments that objective-chance evidence does not involve propensity reasoning effectively claim that the evidence is offered simply to demonstrate the first step of the model above—the rejection of the null hypothesis that $\alpha_* = 0, \alpha_1 = 0, \alpha_2 = 0, \ldots,$ and $\alpha_{n-1} = 0$ in favor of the alternative hypothesis that $\alpha_* = 1, \alpha_1 = 1, \alpha_2 = 1, \ldots,$ *or* $\alpha_{n-1} = 1$, that at least one of the events is in line with what the offering party is attempting to prove. For example, they claim that the evidence of the defendant in *Rex v. Smith* having two subsequent

204. *Id.*

wives who drowned is relevant without propensity reasoning simply to show that *at least one* of the three drownings was not accidental.[205]

Let us consider two modes of reasoning that could follow. First, it is often implied that the chain of reasoning stops here. But this reasoning does not go far enough to be relevant. Rather, the rejection of the null hypothesis that all drownings were accidental must then be connected to the event at issue. The obvious connection is in line with the analysis above: because it is highly unlikely that all three drownings were accidental, a factfinder should replace the null hypothesis of accidental drownings, or randomness, with one that reflects the defendant's propensity to commit murder. In other words, there is a very low likelihood that all drownings were due to accident; therefore, at least one was by the defendant's design; and therefore, the defendant is likely to have acted in accordance with that design and murdered his wife. This much is necessary to relate the rejection of the null hypothesis to the event at issue. Without it, we can only say, "One or more of the drownings was not accidental," when in fact we need to say, "At least one or more of the drownings was not accidental, *and therefore the drowning at issue is likely to be by design rather than accidental*." This is propensity reasoning.

Second, in the alternative, it could be argued that the conclusion that at least one of the drowning events was not accidental itself increases the likelihood that the event at issue—*only as a member of the class of drowning events*—was by design rather than accidental. In other words, it is very unlikely that all of the murders were due to accident; therefore, it is likely that at least one was by the defendant's design; and therefore, because the event at issue is one of three drowning events for which there is a likelihood that at least one was by the defendant's design, it shares one-third in that increased probability.[206]

It is likely that the probative value of objective-chance evidence arises from both of these inference chains. However, the problem with the second chain of reasoning is that it is entirely overshadowed by the former chain of reasoning. In particular, the probative value associated with this reasoning is altogether eclipsed by the probative value of the former reasoning—and by the prejudice caused by the former reasoning, if impermissible.[207] If a prosecutor in an arson case seeks to introduce evidence that the defendant's past five houses burned down, each time following his purchase of insurance, it is possible to make the inference that it is highly unlikely that all of these events were accidental; that, therefore, at least one of the six fire events was by the defendant's design; and that, therefore, the "at least one" non-accidental event may have been the event

205. *See supra* notes 113–125 and accompanying text.

206. *See generally supra* notes 114-119 and accompanying text.

207. *See supra* Subpart I.D. *See generally* Morris, *supra* note 121, at 200 n.74 (highlighting the similarity between propensity reasoning and the purported non-propensity reasoning in Professor Imwinkelried's logic and arguing that "[e]ven if Professor Imwinkelried were correct," the "evidence used on this theory would often fall to the Rule 403 balancing test").

at issue in the case. But the far more obvious chain of inferences—and, arguably, the unavoidable chain of inferences from a behavioral perspective—is that the defendant has a propensity to burn down his houses to collect insurance money.[208]

Moreover, relying on the second chain of reasoning for admissibility superficially covers up the real issue: even if this evidence—say, regarding drowning events of two subsequent spouses—indeed requires character-based propensity reasoning, it seems different from other forms of character evidence. For example, it seems far more probative than ordinary character evidence—as though not allowing this evidence hides the crux of the case from the jury. It begs the question, "Is there something different about this propensity reasoning that makes it categorically more probative?" After all, it is the first chain of reasoning above that is extremely probative and that, putting the ban on character evidence aside, seems to make the most sense to apply. Perhaps propensity reasoning *should* be permitted for objective-chance evidence.

In sum, objective-chance evidence is unlikely to be admissible based on the second chain of reasoning. The probative value of this reasoning is eclipsed by that of the first chain of reasoning. At the very least, it is likely to be rejected based on Rule 403, given the degree to which it is overshadowed by the probative value of the first chain, involving character-based propensity reasoning. In any event, the first chain of inferences, involving propensity reasoning, is what the offering party is really after. It is what really gives objective-chance evidence its whopping probative value.

B. An Exception for Objective-Chance Evidence

Rule 404 is generally understood as a particular application of Rule 403.[209] There are numerous exceptions to the rule against character evidence that replace Rule 404 with a Rule 403 balancing.[210] The substantial problems emphasized in this Article do not arise from formal exceptions. Rather, they arise from the courts' tendency to misread Rule 404(b)(2), misapply the doctrine of chances, carve out ad hoc exceptions, and ignore Rule 404 altogether.[211]

The discussion in Part II provides an explanation for judicial departures from Rule 404. However, it also provides a justification for an exception to Rule 404 for objective-chance evidence.

The rule against character evidence reflects the general tendency of character evidence to involve a risk of unfair prejudice that substantially

208. Indeed, the chain of inferences that relies on propensity reasoning seems necessary in many applications of the doctrine of chances in the courtroom.

209. *See supra* note 42 and accompanying text.

210. *See, e.g.*, FED. R. EVID. 404(a)(2)(A)–(C), 404(a)(3), 413–15.

211. *See supra* Part I.

outweighs the probative value of the evidence.[212] In terms of the aggregation-evidence model, this is in part due to the law's particular aversion to bias arising from character-based inferences—a form of policy-disfavored bias.[213] As with many rules that replace balancing tests, there are costs at the margins associated with the sacrifice of case-specific analysis. However, objective-chance evidence is arguably different. It does not simply involve accuracy costs at the margins of the rule against character evidence. Rather, as discussed in Part II, the very nature of objective-chance evidence—properly defined—sets it apart from other types of character evidence. Categorically, accounting for both precision and bias—forms of probative value and unfair prejudice—and even accounting for the law's particular aversion to character-based bias, it involves a far greater potential to improve accuracy than other forms of character evidence.[214]

In turn, the very substantial accuracy benefits associated with objective-chance evidence place immense pressure on courts to depart from the rule against character evidence. Courts carve out exceptions to the rule or ignore it altogether when the rule does not conform to an intuitive sense of admissibility.[215] This occurs frequently in circumstances involving objective-chance evidence.[216] As discussed above, it can often seem counterintuitive or even unreasonable to exclude this type of evidence.[217] For example, evidence regarding the drowning deaths of the defendant's subsequent wives in *Smith* seems central to an accurate judgment; evidence regarding five prior drug-trafficking arrests seems central to proving that a drug-trafficking defendant had knowledge of the presence of drugs in their automobile. Therefore, courts frequently depart from the rule against character evidence to admit objective-chance evidence, often without even acknowledging the rule or perhaps without realizing the rule's application.[218]

Indeed, many of the exceptions and departures discussed in Subparts I.B and I.C above involve objective-chance evidence. For example, anecdotal evidence to prove discriminatory intent in discrimination cases, prior-crime evidence to prove knowledge or intent in drug-trafficking cases, prior-crime evidence to prove intent in fraud cases, and various forms of statistical evidence all frequently involve objective-chance evidence.[219]

In cases in which courts address the admission of objective-chance evidence notwithstanding Rule 404, they frequently explain it based on (1) the permissibility of the evidence for a non-propensity purpose under the doctrine

212. *See supra* Subpart I.A.
213. *See supra* Subpart II.C.
214. *See supra* Subparts II.C–E.
215. *See supra* notes 201–204 and accompanying text.
216. *See supra* Subparts I.B–D.
217. *See supra* Subpart II.E.
218. *See supra* Subparts I.B–D.
219. *See supra* Subparts I.B–C.

of chances or (2) the permissibility of the evidence as an exception to the rule against character evidence under Rule 404(b)(2).[220] However, both of these bases are problematic: they propagate misreadings of Rule 404(b)(2) as providing for exceptions to the rule against character evidence;[221] they encourage common-law carve-outs from Rule 404;[222] and they promote misapplications of the doctrine of chances.[223] Moreover, such admissions give rise to substantial unpredictability by facilitating evidentiary decisions that are based on intuition and then justified using vague rules and poor precedent, or not justified at all.[224]

Relatedly, the courts' frequent understanding of the doctrine of chances as involving a non-propensity use of other-acts evidence has promoted *missed* applications of the doctrine and has limited its usefulness. Specifically, this understanding has prevented courts from applying the doctrine of chances in circumstances in which it could clearly apply—and in which the doctrine would be very useful to explain an admissibility decision—if it were understood as an *exception* to the rule against character evidence. Instead, courts rely on avoidance and loose reasoning to admit evidence that seems essential for an accurate verdict but that clearly involves character-propensity reasoning.

For example, as discussed in Part I, courts frequently rely on a misreading of Rule 404(b)(2) as an exception to the rule against character evidence in order to admit objective-chance evidence to prove absence of mistake or accident. This misapplication of Rule 404(b)(2) creates or exacerbates poor precedent for future cases. It propagates an overbroad exception that covers all of the purposes referred to in Rule 404(b)(2)—an exception that altogether eats up the rule against character evidence—rather than an exception for objective-chance evidence in particular. A future court may then cite this precedent as a basis for admissibility even for ordinary character evidence. However, if the courts adopted an explicit objective-chance exception to the rule against character evidence, then this evidence could be correctly classified as character evidence and then admitted under such an exception. Indeed, various categories of Rule 404 departures discussed in Subparts I.B and I.C—departures that currently cause confusion, inconsistency, and disarray, such as those for

220. *See supra* Subpart I.B.

221. *See supra* Subparts I.B–C.

222. *See supra* Subpart I.B–C.

223. *See supra* Subpart I.D; *see also* People v. Burnett, 2 Cal. Rptr. 3d 120, 130–32 (Ct. App. 2003) (examining the admissibility of evidence that the defendant "beat a stray dog to death" to rebut the defendant's accident defense); FISHER, *supra* note 19, at 194.

224. *See supra* Subpart I.B; *see also* Goode, *supra* note 202, at 718, 802–10 (arguing that "decision-making under Rule 404(b) can be improved only by making it easier for courts to admit high-probative-value other-acts evidence without having to engage in character-propensity denial," including amending Rule 404(b) to include a "true exception" for proof of "a criminal defendant's intent through a character-propensity inference unless the defendant agrees not to controvert state of mind.").

anecdotal evidence in discrimination cases and prior-crimes evidence in drug trafficking cases—would likely fall within an objective-chance exception.

In this Article, I have defined objective-chance evidence rigorously and broadly. I have shown that it is best understood as a form of character evidence. An explicit objective-chance exception to the rule against character evidence—one that replaces Rule 404 with a rigorous Rule 403 balancing for objective-chance evidence—would account for the unique accuracy-enhancing nature of objective-chance evidence and permit a more coherent rule against character evidence that courts could more strictly adhere to.

Such an exception would allow courts to address the accuracy benefits of objective-chance evidence head-on and would allow admissibility decisions based on those benefits—i.e., based on a balancing of probative value and unfair prejudice—rather than on misreadings of Rule 404(b)(2) and weak logic surrounding the applicability or inapplicability of Rule 404.[225] It would allow flexibility for courts to consider the unique nature of objective-chance evidence while also considering the law's strong aversion to character-based bias.[226]

Below, I discuss a number of important considerations surrounding an objective-chance exception to the rule against character evidence. Specifically, I discuss (1) considerations regarding a Rule 403 balancing for objective-chance evidence; (2) whether clear lines can be drawn between objective-chance evidence and other forms of character evidence such that an objective-chance exception will enable a more coherent rule against character evidence; (3) why objective-chance evidence is distinct from other forms of character evidence in that it avoids promoting juror reliance on implicit biases; and (4) how courts can institute an objective-chance exception to the rule against character evidence.

C. A Rigorous Rule 403 Balancing

Although objective-chance evidence involves assurances of probative value far more substantial than those associated with character evidence generally—thus arguably justifying an exception to the general ban on character evidence—it can also be very prejudicial. The rationales underlying Rule 404—for example, the risk of a jury overweighting character evidence or punishing an individual for acts not at issue in the case—pose substantial dangers to the fairness of a trial, even if they do not justify a wholesale exclusion of objective-chance evidence. Therefore, a Rule 403 balancing is not just necessary; rather, it should

225. Arguably, certain existing exceptions to the rule against character evidence can also be explained based on the aggregation-evidence model discussed in Part II—for example, the various standards applied to impeachment evidence, exceptions for character evidence in sexual assault cases, and standards surrounding good-character evidence. *See* FED. R. EVID. 404(a)(2)(A)–(C), 413–15, 608–09.

226. *See infra* Subpart III.C.

be conducted with particular sensitivity to the general ban on character evidence and the law's aversion to character-based bias.[227]

I am not suggesting a balancing test distinct from Rule 403. Rather, Rule 403's balancing test is well-equipped to account for the particular concerns surrounding objective-chance evidence. In conducting this balancing, courts should consider the bias–variance trade-off and a number of important principles associated with aggregation evidence. First, bias is inherent in other-acts evidence, and it is, in a sense, necessary to achieve the probative value associated with the evidence.[228] Second, the evidence will be most beneficial when judgment variability (in the absence of the objective-chance evidence) is high and uncertainty surrounding the objective-chance evidence is low.[229] As discussed, the very nature of objective-chance evidence provides certain assurances that these criteria will be met relative to ordinary character evidence. But these features will vary from case to case. Third, the comparability of the other-acts evidence will minimize the bias that it causes in some respects but can cause bias in other respects; comparability should therefore receive careful attention in determining admissibility under Rule 403.[230] Finally, courts should consider the factors above, and others, in light of the particular aversion that the law has toward character-based bias. Bias is weighted quite heavily, and although Rule 403 generally favors admissibility, objective-chance evidence is a form of other-acts character evidence, and it requires substantial benefit in terms of variance-reduction, or probative value, to justify admissibility. Objective-chance evidence involves an unusual degree of potential to yield such value. But a court should only admit it if it satisfies a rigorous Rule 403 balancing.

D. Enabling a Strict Adherence to Rule 404 by Discerning Questions of Chance

If courts apply a broad, well-defined objective-chance exception combined with a rigorous Rule 403 balancing, there will be substantially less pressure for courts to chip away at Rule 404. A primary aim of the proposed approach is to facilitate an environment in which courts do not need to choose between, on the one hand, adhering to Rule 404 and making unreasonable evidentiary exclusions and, on the other hand, violating Rule 404 to allow admission of unusually probative evidence.

As suggested above, courts frequently apply the doctrine of chances in particular types of cases as a non-propensity use of other-acts evidence while ignoring or avoiding it in a wide range of contexts in which the doctrine would

227. *See generally* United States v. Guardia, 135 F.3d 1326, 1330–32 (10th Cir. 1998) (discussing Rule 403 balancing in sexual assault cases).

228. *See generally An Aggregation Theory of Character Evidence*, *supra* note 17, at 48–49.

229. *See supra* Subpart II.E.

230. *See supra* Subpart II.C.

apply were it defined explicitly and properly as an exception to the rule against character evidence.[231] By formulating a well-defined and appropriate exception, my aim is for courts to better be able to adhere strictly to the rule against character evidence. Thus, this Article argues that applying Rule 403, rather than a categorical exclusion, to determine the admissibility of objective-chance evidence may be fundamental to an otherwise strict adherence to Rule 404. Similarly, correctly characterizing objective-chance evidence as an exception to Rule 404, rather than as evidence that does not depend on propensity reasoning in the first instance, is an important step toward a logical and well-reasoned approach to its admissibility.

The proposed method therefore aims to reduce uncertainty and arbitrariness surrounding the admissibility of character evidence. It aims to accomplish this by applying a logical and well-defined exception that accounts for the areas in which character evidence is unusually probative and in which courts have the most pressure to ignore or reject the rule against character evidence. Additionally, the proposed method promotes a rigorous Rule 403 analysis that is particularly sensitive to bias caused by character-based propensity reasoning.

However, excepting objective-chance evidence from the ban on character evidence involves an important risk—a risk not only of watering down the sought-after effect but one of altogether exacerbating the uncertainty surrounding character evidence.[232] It risks allowing the exception to swallow the rule. Therefore, in order to avoid a situation in which an exception for objective-chance evidence is used to further deteriorate the ban on character evidence, we must consider whether litigants can inappropriately reframe ordinary character evidence as objective-chance evidence.

Consider a robbery case in which the prosecutor wishes to introduce evidence that the defendant has committed robberies in the past for the purpose of making the inference that the defendant has a character for committing robbery and is therefore more likely to have acted in accordance with that character and committed the robbery in question. This is blatant impermissible character evidence under Rule 404.[233] But assume now that courts admit character evidence if it qualifies for the objective-chance exception, as defined above. Knowing that this evidence would make their case, the prosecutor attempts to reframe the evidence—ordinary character evidence under Rule 404—as objective-chance evidence. Rather than rely on the foregoing ordinary propensity reasoning, the prosecutor argues that the evidence should be admitted for the following chain of inferences: there have been multiple instances in which the defendant has been suspected of robbery;

231. *See supra* Subpart I.D.

232. *See generally An Aggregation Theory of Character Evidence*, *supra* note 17, at 59.

233. *See* FED. R. EVID. 404(b)(1).

there is a very low likelihood that the defendant would fall under suspicion this many times by chance if the defendant had nothing to do with the robberies; therefore, the defendant is highly likely to have committed robbery in at least one of these instances; therefore, the defendant is likely to have committed robbery in the case at hand.[234]

There are two important points to take away from this scenario. First, it is easy to superficially reframe character evidence as objective-chance evidence. Second, however, the reframing is only superficial; behind the reframing, there are important differences between ordinary character evidence and the special category of objective-chance evidence. It is critical for courts to distinguish genuine objective-chance evidence from ordinary character evidence repackaged to look like objective-chance evidence.

Compare the robbery scenario above to contexts in which we have discussed genuine objective-chance evidence. Many of these contexts have involved showings of mental state. The question is, for example, whether an event just happened by chance or, in the alternative, there was in fact intent, motive, or knowledge. For example, was the defendant simply in the wrong car at the wrong time, or did the defendant *know* that the car was being used to transport drugs? Even when these cases involve a question of conduct, *genuine* objective-chance evidence rarely involves questions of identity—of whether *this* is the person who committed the crime. For example, in the bathtub case, *Rex v. Smith*, the question is not *who* murdered Mrs. Smith, but *whether* Mrs. Smith was murdered.[235] The question is whether the drowning was an accident or by the defendant's design.

This comparison suggests the fundamental distinction between objective-chance evidence and ordinary character evidence, such as the evidence offered in the robbery example above. Objective-chance evidence always involves the following question: Is this event due to chance or to something else—i.e., to the hypothesized element, such as design, intent, or knowledge? For example, is the fact that the defendant was in a car used for drug trafficking due to chance—i.e., bad luck—or did the defendant have the knowledge or intent required for a conviction? Other-acts evidence may well be probative of this fact issue. It qualifies for the objective-chance exception, as defined above. Similarly, in the bathtub case, the question is whether the defendant's wife drowned by chance or by the design of the defendant. Note that "by chance" is with reference to the defendant in particular. It does not matter, for purposes of our discussion, whether the drowning was literally accidental or, for example, by suicide. Both qualify for the drowning being "by chance" with respect to the defendant. What matters, in terms of the defendant's guilt or innocence, is

234. *See generally* Imwinkelried, *supra* note 8, at 11–12 (discussing the "recent extension [of the doctrine of chances] to prove the accused's identity as the perpetrator").

235. Rex v. Smith, (1915) 84 LJKB 2153.

whether the drowning was by his *design* or just due to another cause that was out of his intent or control.

This fact question of chance versus an alternative hypothesis is not a material issue in the robbery example. There is, for example, no question of accident, intent, or design. We know that the robbery occurred and that there was intent by whoever committed the crime. What we do not know is who did it. The question of chance posed by the prosecutor in that example—whether it is chance that the defendant has fallen under suspicion of robbery so many times—is *not* a material issue in the case. It is probative only through ordinary propensity reasoning—injecting prior acts to make the inference that the defendant *was the one who did it*. In contrast, in the bathtub case, the question of chance (versus design) is a material issue: Bessie's drowning raises the fact question, "Was the drowning an accident (i.e., by chance) or by design?"

To qualify for the objective-chance exception, the offering party must, implicitly or explicitly, be able to formulate, in effect, a hypothesis test in which the null hypothesis is that the event at issue occurred by chance or randomness and the alternative hypothesis is that the event at issue occurred by some hypothesized element, such as design or intent.[236] As another example, consider a discrimination case in which the plaintiff, a highly qualified female employee who is passed over for promotion, offers evidence of other events in which the employer passed over highly qualified female employees in favor of less-qualified male employees. This evidence may qualify for the objective-chance exception since it is offered to prove that the employer's decision to promote a particular less-qualified male employee rather than the plaintiff was not due to chance (given the features of the two employees) but was rather due to the design, or the illegitimate purpose, of the employer. Again, this is different than the robbery case, in which there is no issue of chance versus an alternative hypothesis of, e.g., design. The evidence in the robbery example is simply offered to show, through ordinary propensity reasoning, that it was the defendant who committed the crime.[237]

---

236. In this Subpart, I show how courts can distinguish objective-chance evidence from ordinary character evidence based on the nature of the evidence and the inferences that it involves. However, objective-chance evidence can also be distinguished explicitly in terms of the aggregation-evidence model described in Part II. Specifically, the robbery example—whether reframed superficially as objective-chance evidence or not—does not involve a particularly high level of judgment variability. It simply involves the common question, "Who did it?"—not the uniquely unpredictable fact determinations that objective-chance evidence generally informs. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 58–61. Moreover, it does not involve the type of precise information that genuine objective-chance evidence involves. Even as reframed, its informational value is only based on ordinary character inferences—not the extremely precise signal that objective-chance evidence involves with respect to the issue of design versus accident. *Id.* In close cases, a court can look to these underlying concerns to determine whether certain evidence constitutes objective-chance evidence.

237. To make this conclusion more obvious, construct a hypothetical that limits the uncertainty surrounding the other-acts evidence. Assume, for example, that in the robbery illustration, there is overwhelming evidence that the defendant committed the prior acts of robbery rather than only enough to suspect the defendant. Ordinarily, reducing the uncertainty surrounding the other-acts evidence should

This makes sense as a requirement because, after all, the distinguishing feature of objective-chance evidence in the formal definition provided above, and in existing definitions in the case law and literature, is that multiple other events of the same general kind demonstrate that the event at issue was not due to chance or accident.[238] The proposal herein allows for the use of objective-chance evidence as an exception to the ban on character evidence; however, the evidence must be probative of whether an event at issue occurred by chance or by some other hypothesized force.

E. Avoiding Dependence on Prior Beliefs and Prejudices

One important criticism of the exceptions to the rule against character evidence is that such exceptions invite jurors to rely heavily on their implicit biases, or prior beliefs and prejudices, regarding a defendant's (or other litigation party's) race, appearance, and other background characteristics in making character inferences and determining a verdict.[239] However, unlike other forms of character evidence, jurors are unlikely to rely heavily on prior beliefs and prejudices when making inferences based on objective-chance evidence.

In a separate Article, I have applied the aggregation-evidence model to argue that character evidence requires jurors to rely on their preexisting beliefs, prejudices, and stereotypes to make inferences regarding an individual's character and whether the individual acted in accordance therewith.[240] To summarize: Jurors begin to form "priors" regarding a defendant's character the moment they sit for trial. They do this by applying their preexisting beliefs, prejudices, and stereotypes to the background characteristics of the defendant—for example, the defendant's race, age, economic status, accent, etc.[241] These priors remain relatively subdued because they provide a very imprecise informational signal relative to the evidence in the case (and especially when the court emphasizes the importance of deciding a verdict based on the evidence). However, when the court admits character evidence, it invites jurors

provide *more* information and thus improve the argument for admissibility. However, in this example, the opposite occurs. It becomes clear that the argument boils down to basic character evidence: the defendant committed robberies in the past and is therefore more likely to have committed the robbery in question. The fake question of chance posed by the prosecutor disappears. In the bathtub case, on the other hand, this mental exercise takes us in the opposite direction. If we assume that there is overwhelming evidence that the defendant in fact murdered two subsequent wives, the question of chance as to the act at issue remains: the crux of the case is still the question whether Mrs. Smith's drowning was accidental or by design. Now, however, as we would expect, the evidence becomes even more probative of the question of chance versus design since we know with relative certainty that the defendant in fact drowned his subsequent wives.

238. *See An Aggregation Theory of Character Evidence*, *supra* note 17, 55–57.

239. *See* Hillel J. Bavli, *Character Evidence as a Conduit for Implicit Bias*, 56 U.C. DAVIS L. REV. (forthcoming 2023) (on file with author).

240. *Id.* (manuscript at 3–8).

241. *Id.* (manuscript at 26–33).

to rely heavily on their priors by (1) asking jurors to assess precisely what their priors speak to—the defendant's character—and (2) providing very little information to guide the jurors' character assessments.[242] For example, jurors are generally provided with a very small and biased sample of prior acts and no information regarding the strength or consistency of the defendant's propensity or the consistency of people's behavior more generally.[243] Consequently, jurors fill substantial informational gaps regarding the defendant's character with their priors—which relate precisely to the defendant's character.[244]

For example, consider a hypothetical assault case in which the court admits two prior incidents in which the defendant is alleged to have committed assault. From this evidence, jurors must judge the defendant's character for violence and make inferences regarding the likelihood with which the defendant acted in accordance with the assessed character for violence on the occasion in question. The jurors have very little information to guide these inferences. Therefore, they fill substantial evidentiary gaps with their priors regarding the defendant's race, sex, economic status, and other background characteristics.[245]

However, a court's admission of objective-chance evidence is unlikely to invite a juror's reliance on their prior beliefs as it would with the admission of other forms of character evidence. This is due to the same factors that make objective-chance evidence so accuracy-enhancing in the first instance: the level of precision of the already-existing evidence and the level of precision of the priors. Specifically, objective-chance evidence involves a highly precise informational signal regarding the question at issue—chance versus design.[246] The precision with which objective-chance evidence speaks to the question of chance versus design, in a sense, crowds out a juror's application of priors, which are highly imprecise for the question at hand.

Remember, pursuant to the aggregation-evidence model, jurors rely most heavily on this type of evidence when the aggregation evidence is precise relative to the already-existing evidence.[247] If there is already abundant and precise evidence directly related to the event at issue, then aggregation evidence adds little value—especially if the aggregation evidence is highly imprecise relative to the other evidence. This is the case with prior beliefs relative to objective-chance evidence (as opposed to prior beliefs relative to other forms of character evidence and as opposed to objective-chance evidence relative to the non-character evidence in a case). The issue is whether the act in question is due to chance or design. This question is answered precisely by the objective-chance evidence and imprecisely by a juror's prior beliefs.

242. *Id.*
243. *Id.* (manuscript at 30–32).
244. *Id.*
245. *See id.*
246. *An Aggregation Theory of Character Evidence*, *supra* note 17, at 60.
247. *See supra* Part II.

Relatedly, although objective-chance evidence leads to a chain of inferences that employs propensity reasoning, it involves reasoning that is distinct from that associated with other forms of character evidence. Specifically, character evidence invites jurors to rely on their priors in part because of the symmetry between the priors and the question at hand.[248] After all, the priors speak precisely to what the juror is being asked to assess—the defendant's character. With objective-chance evidence, on the other hand, jurors are not asked to assess prior acts to make inferences directly regarding the defendant's character. Rather, jurors are asked to consider the likelihood of the type of event in question and to reject the hypothesis of "accident" in favor of one of "design" if, in light of the number of similar events evidenced, the jurors conclude that the events could not have occurred randomly. For this question, the juror's priors are relatively imprecise.

In summary, although there is a strong argument that a verdict based on character evidence will generally be highly influenced by the jurors' prior beliefs and prejudices regarding the background characteristics of the defendant, the same factors that make objective-chance evidence so accuracy-enhancing also suggest that, contrary to other forms of character evidence, jurors will rely most heavily on this evidence and not on their prior beliefs and prejudices.

F. Instituting an Objective-Chance Exception

Rule 404 requires Advisory Committee clarification to be applied uniformly throughout the federal courts. As discussed above, courts vary wildly in how they interpret Rule 404. They "have grown increasingly permissive in allowing the admission of other-acts evidence," and they "routinely admit the previous uncharged misdeeds of criminal defendants, threatening to undermine the bedrock ban on character evidence."[249]

However, immediate change is necessary. Some courts have recognized this, and as Professors Capra and Richter have emphasized, "There is a war raging over the admissibility of the prior bad acts of criminal defendants in federal trials."[250] Central to the disarray surrounding the rule against character evidence is the lack of clarity regarding Rule 404. However, courts arguably should not wait for additional guidance. As reflected in the wide variety of interpretations of and approaches to Rule 404 in the courts, there may well be sufficient flexibility in Rule 404 for courts to implement change immediately.

Indeed, the Advisory Committee has arguably signaled such flexibility with respect to a possible objective-chance exception. In response to recommendations for substantive reforms to Rule 404(b)—and, in particular,

248. Bavli, *supra* note 239 (manuscript at 30–32).

249. Capra & Richter, *supra* note 51, at 778.

250. *Id.* at 769.

recommendations for "criteria for [a] more careful application" of the Rule, and one that protects an exclusionary interpretation of Rule 404(b) and limits the ability of courts to carve out other-acts exceptions to it—the Advisory Committee "determined that it would not propose substantive amendments to Rule 404(b), because they would make the Rule more complex without rendering substantial improvement."[251] Among other things, the Committee explicitly concluded:

> [A]n attempt to require the court to establish the probative value of a bad act by a chain of inferences that did not involve propensity would add substantial complexity, while ignoring that in some cases, a bad act is legitimately offered for a proper purpose but is nonetheless bound up with a propensity inference . . . .[252]

As an example, the Committee referred specifically to the doctrine of chances, used "to prove the unlikelihood that two unusual acts could have both been accidental."[253]

Arguably, therefore, the courts would be well within their authority to adopt an approach to character evidence that is in line with the method proposed herein.

CONCLUSION

This Article addresses an important problem in evidence law: the state of disarray surrounding other-acts character evidence and the arbitrariness in case outcomes that results from it. I have argued that a significant source of the problem is that, while most character evidence would likely fail a Rule 403 balancing test based on the law's aversion to character-based bias, a certain category of character evidence—objective-chance evidence—is generally accuracy-enhancing, even accounting for the heavy weight of character-based bias. This category of evidence has placed substantial pressure on courts to carve out exceptions to Rule 404 or to ignore it altogether. This effect is only exacerbated by a lack of clarity regarding the meaning of Rule 404(b) and by confusion regarding the boundaries of and central idea behind the doctrine of chances.

In this Article, I apply the aggregation-evidence model to address long-standing disagreement and confusion over whether the doctrine of chances relies on character-based propensity reasoning. I argue that it does, at least when applying the chain of inferences necessary to achieve the primary and

---

251. ADVISORY COMM. ON EVIDENCE RULES, JUD. CONF. OF THE U.S., REPORT OF THE ADVISORY COMMITTEE ON EVIDENCE RULES 4–5 (May 2018) [hereinafter ADVISORY COMMITTEE REPORT]; *see* United States v. Thorne, No. 18-389, 2020 WL 122985, at *5 n.4 (D.D.C. Jan. 10, 2020) (noting that the Committee "recently considered but ultimately rejected substantive changes to Rule 404(b)").

252. ADVISORY COMMITTEE REPORT, *supra* note 251, at 5; *see Thorne*, 2020 WL 122985, at *5 n.4.

253. ADVISORY COMMITTEE REPORT, *supra* note 251, at 5.

overwhelming probative value of objective-chance evidence. I then use this model to formulate a broad but well-defined objective-chance exception to the rule against character evidence—one that is based on scientific principles of aggregation and accuracy and that accounts for many of the courts' departures from Rule 404. I show why this exception, combined with a rigorous Rule 403 balancing, may lead to a stricter adherence to Rule 404.

The proposed method may improve the accuracy of judgments while promoting a more consistent and principled approach to character evidence. It may also facilitate a trial that better comports with constitutional values, norms of fairness, and principles of equality.