



FLORIDA LAW REVIEW

Volume 70, Page 347
Published 2018

Convictions Based on Character: An Empirical Test of Other-Acts Evidence

Michael D. Cicchini & Lawrence T. White

Recommended citation:

Michael D. Cicchini & Lawrence T. White, *Convictions Based on Character: An Empirical Test of Other-Acts Evidence*, 70 FLA. L. REV. 347 (2018)

This article is also available at: <https://scholarship.law.ufl.edu/flr/vol70/iss2/2/>

Other works by Michael D. Cicchini are available at: <https://cicchinilaw.com/articles>

CONVICTIONS BASED ON CHARACTER: AN EMPIRICAL TEST OF OTHER-ACTS EVIDENCE

*Michael D. Cicchini**

*Lawrence T. White***

“So my word is not enough; my promise worthless;
the fact that I have served my time nothing but the
emblem of my continuing guilt.”¹

Abstract

Despite the time-honored judicial principle that “we try cases, rather than persons,” courts routinely allow prosecutors to use defendants’ prior, unrelated bad acts at trial. Courts acknowledge that jurors could improperly use this other-acts evidence as proof of the defendant’s bad character. However, courts theorize that if the other acts are also relevant for a permissible purpose—such as proving the defendant’s identity as the perpetrator of the charged crime—then a cautionary instruction will cure the problem, and any prejudice is “presumed erased from the jury’s mind.”

We put this judicial assumption to an empirical test. We recruited 249 participants to serve as mock jurors in a hypothetical criminal case. After reading the identical case summary, jurors were randomly assigned to one of two groups, each of which received different evidence on the issue of identity. Group A received conclusive proof, in the form of a stipulation, that if a crime was committed, the defendant was the one who committed it. Group A convicted at the rate of 33.1%. Group B received less certain evidence of identity in the form of the defendant’s somewhat similar, prior conviction, along with a cautionary instruction that this other act may not be used as evidence of the defendant’s character. Group B convicted at the much higher rate of 48.0%.

The difference in conviction rates is statistically significant. Further, jurors in Group B were also more confident in their verdicts despite receiving less certain evidence of guilt and a cautionary instruction. These empirical findings demonstrate that cautionary instructions are not

* J.D., *summa cum laude*, Marquette University Law School (1999); C.P.A., University of Illinois Board of Examiners (1997); M.B.A., Marquette University Graduate School (1994); B.S., University of Wisconsin—Parkside (1990). Michael Cicchini is a criminal defense lawyer in Kenosha, Wisconsin. He is also the author of several books and law review articles on criminal and constitutional law.

** Ph.D., University of California, Santa Cruz (1984); M.A., *with distinction*, California State University at Fresno (1979); B.A., *with honors*, Whittier College (1975). Dr. White is Professor of Psychology at Beloit College. He also directs the College’s Law & Justice Program.

1. HENRY FOOL (Sony Pictures Classics 1998).

effective, and jurors will use other-acts evidence for impermissible purposes including, for example, the forbidden character inference. Given this, we discuss several pretrial strategies for defense counsel to limit the prejudicial impact of other-acts evidence.

- INTRODUCTION348
 - I. “WE TRY CASES, RATHER THAN PERSONS”350
 - II. OTHER-ACTS EVIDENCE352
 - III. CAUTIONARY INSTRUCTIONS355
 - IV. THE STUDY357
 - A. *Hypothesis*357
 - B. *Participants*357
 - C. *Study Design*358
 - D. *Identity Evidence*358
 - E. *Findings*362
 - V. DISCUSSION: THE IMPACT OF OTHER ACTS364
 - VI. PRETRIAL STRATEGIES FOR THE DEFENSE365
 - A. *Relevance*365
 - B. *Unfair Prejudice*367
 - C. *The Stipulation*369
 - VII. POTENTIAL CRITICISMS OF THE STUDY372
 - A. *Case Study Method*372
 - B. *Single Fact Pattern*374
 - C. *Lack of Deliberations*374
 - D. *Participant Attention Level*375
 - E. *Participant Bias*375
- CONCLUSION376

INTRODUCTION

The Federal Rules of Evidence generally prohibits the government from introducing evidence at trial of a criminal defendant’s bad character. Instead, to win a conviction, the prosecutor must prove what the defendant actually did with regard to the charged crime. What the

defendant may have done in the past and, more specifically, what kind of person he is should not serve as the basis for a new conviction.²

Despite this time-honored prohibition on character evidence, courts still permit prosecutors to use a defendant's prior, unrelated bad acts at trial. Courts acknowledge that such other-acts evidence goes to the defendant's character; however, as long as the prosecutor offers the other acts ostensibly for a permissible purpose—such as proof of the defendant's identity, intent, or absence of accident with regard to the charged crime—courts typically allow the prosecutor to use the evidence.³

Due to the highly prejudicial nature of other-acts evidence, courts instruct jurors that they are not to use the defendant's other acts to decide that he is a bad person and is therefore guilty of the charged crime. Rather, jurors are to use the other acts only for a limited purpose—for example, to decide whether the defendant is the perpetrator or whether the defendant's actions were intentional or accidental. Once a court gives this cautionary instruction, all unfair prejudice is presumed to be wiped from jurors' minds.⁴

Given the obvious incompatibility between this judicial assumption and the way the human mind actually works, we decided to conduct a controlled study on other-acts evidence, cautionary instructions, and the issue of identity. We recruited 249 participants to serve as mock jurors in a hypothetical criminal case. After reading a case summary—including the elements of the crime, a summary of trial testimony, and an instruction on the burden of proof—participants were randomly assigned to one of two groups, each of which received different evidence on the issue of identity.⁵

Group A (N = 124) received conclusive proof on the issue of identity in the form of a stipulation and convicted the defendant at a rate of 33.1%. Group B (N = 125) received far less certain evidence on the issue of identity: the defendant's somewhat similar, prior bad act, along with a cautionary instruction that such evidence should not be used to decide that the defendant has a bad character and is therefore guilty of the charged crime. Even though the defendant's other act was inconclusive evidence of identity and was accompanied by a cautionary instruction, Group B convicted the defendant at a rate of 48.0%.⁶

Group B's conviction rate was nearly 50% higher than Group A's, a statistically significant difference. Jurors in Group B were also more certain they had chosen the correct verdict. These empirical findings

2. See *infra* Part I.

3. See *infra* Part II.

4. See *infra* Part III.

5. See *infra* Sections IV.A–IV.D.

6. See *infra* Section IV.E.

provide strong evidence that the courts' assumption is wrong. The evidence demonstrates that cautionary instructions are not effective in limiting the prejudicial impact of other-acts evidence. Rather, jurors use other acts for impermissible purposes, such as determining that the defendant has a bad character and, for that reason, is guilty of the crime charged.⁷

Part I of this Article discusses the policy behind the prohibition on character evidence. Part II discusses the rule on other-acts evidence and explains how prosecutors circumvent the character-evidence prohibition by ostensibly using the other acts for permissible purposes. Part III then discusses, and gives examples of, the cautionary instructions that are intended to protect the defendant from the impermissible character inference associated with other acts.

Part IV of this Article describes our controlled study, including our hypothesis, our study design, and our empirical findings. Part V discusses the findings in more detail and further explores the prejudicial impact of other-acts evidence. Part VI then offers pretrial strategies for defense counsel when dealing with other-acts evidence. These strategies include the use of our empirical findings to demonstrate that the unfair prejudice of other acts substantially outweighs any probative value. Finally, Part VII discusses the limitations of our study design and considers possible alternative methodologies for other researchers.

I. "WE TRY CASES, RATHER THAN PERSONS"

A deeply rooted principle in criminal law is that a jury should not convict a defendant of a crime because of his prior, unrelated bad acts. Rather, a jury should judge a defendant on the evidence specific to the crime with which he is charged. "[I]n our system of jurisprudence, we try cases, rather than persons, and thus a jury may look only to the evidence of the events in question, not defendants' prior acts in reaching its verdict."⁸

There are many good reasons to cling to this basic tenet of criminal law. If jurors hear evidence of a defendant's other acts, they may convict not because they believe he is guilty of the charged crime, but rather to

7. See *infra* Part V.

8. *People v. Allen*, 420 N.W.2d 499, 504 (Mich. 1988) (citing *United States v. Mitchell*, 2 U.S. (2 Dall.) 348, 357 (1795)). For a history of the rule prohibiting the use of other acts to prove a defendant's character, see Jennifer Y. Schuster, *Uncharged Misconduct Under Rule 404(b): The Admissibility of Inextricably Intertwined Evidence*, 42 U. MIAMI L. REV. 947, 951-58 (1988) (explaining the prohibition "has its roots in Great Britain's Treason Act of 1695").

punish him (again) for his prior misdeeds.⁹ In effect, the defendant would never be able to atone for his other acts. This, of course, would be “antithetical to the precept that ‘a defendant starts his life afresh when he stands before a jury.’”¹⁰

Similarly, jurors may use the other-acts evidence to convict the defendant preemptively, reasoning that, although he may be “innocent momentarily,” society must be protected from the crimes he might commit in the future.¹¹ This desire for the illusion of security is a strong one, and some prosecutors are willing to exploit it in order to win convictions.¹²

Beyond these risks, however, the most significant risk of other-acts evidence is that jurors will “generaliz[e] a defendant’s earlier bad act into *bad character* and tak[e] that as raising the odds that he did the . . . act now charged.”¹³ To illustrate this, assume that a defendant is charged with battery but claims he acted in self-defense. If jurors were to hear evidence of his prior, unrelated battery conviction, they may well find him guilty of the current battery charge—not because they are persuaded by the accuser’s testimony, but because the defendant committed a prior battery. This can cause jurors to conclude that the defendant is a violent or hot-tempered person and is therefore guilty of the charged crime.

Arguably, a defendant’s character could be relevant evidence. After all, most battery crimes are committed by hot-tempered persons; it follows that a defendant’s character trait for being hot-tempered is relevant to show that he “is by propensity a probable perpetrator of the crime.”¹⁴ So the problem is not that “character is irrelevant; on the

9. *See Old Chief v. United States*, 519 U.S. 172, 181 (1997) (quoting *United States v. Moccia*, 681 F.2d 61, 63 (1st Cir. 1982)) (discussing the risk that the jury, “uncertain of guilt . . . will convict anyway because a bad person deserves punishment”).

10. *People v. Crawford*, 582 N.W.2d 785, 790 (Mich. 1998) (quoting *People v. Zackowitz*, 254 N.Y. 192, 197 (1930)).

11. *Old Chief*, 519 U.S. at 180–81.

12. *See, e.g., Hennon v. Cooper*, 109 F.3d 330, 333 (7th Cir. 1997) (finding that “the prosecutor could have been understood to be telling the jury to convict irrespective of Hennon’s guilt or innocence, lest an acquittal be interpreted as a green light for street gangs and drive-by killings”).

13. *Old Chief*, 519 U.S. at 180 (emphasis added).

14. *Michelson v. United States*, 335 U.S. 469, 475 (1948). However, this simplistic reasoning may be flawed. *See Linda S. Eads et al., Getting It Right: The Trial of Sexual Assault and Child Molestation Cases Under Federal Rules of Evidence* 413–15, 18 BEHAV. SCI. & L. 169, 171 (2000) (“[W]e risk inaccuracy when we assume consistency of behavior over time and circumstances . . .”). Further, individuals all around us defy their fundamental character and control their violent urges on a daily basis. State and federal governments even tout this fact to justify their massive expenditures of taxpayer dollars on correctional programs. *See, e.g., NAT’L INST. OF CORRECTIONS, COGNITIVE-BEHAVORIAL TREATMENT: A REVIEW AND DISCUSSION FOR CORRECTIONS PROFESSIONALS* 17 (2007), <http://static.nicic.gov/Library/021657.pdf> (explaining

contrary, it is said to weigh *too much* with the jury, and to so *overpersuade* them as to prejudge one with a bad general record and deny him a fair opportunity to defend against a particular charge.”¹⁵

Given this, the general ban on character evidence is a hallmark of a system that values fundamental fairness and the appearance of fairness. And while it is true that modern rules of evidence have carved into this time-honored principle to some extent,¹⁶ courts still continue to honor, at least superficially,¹⁷ the idea that a defendant should only be convicted “for what he did, not for who he is.”¹⁸ However, even when this character-evidence prohibition remains fully intact, courts still permit prosecutors to use a defendant’s other acts at trial—provided the other acts have some valid purpose in addition to proving the defendant’s character. In fact, as discussed below, the use of other-acts evidence has become incredibly common, if not the norm.

II. OTHER-ACTS EVIDENCE

Most commonly, other-acts evidence takes the form of a defendant’s prior conviction for a similarly-named crime, as in the battery example above. However, prosecutors can use other-acts evidence at a defendant’s trial even if he was found not guilty of the other acts and even if the state previously declined to prosecute the other acts. Further, while other-acts evidence is usually similar to the charged crime, it need not be. And the other acts—regardless of whether they are similar in nature to the current charge—may involve the same alleged victim or someone entirely different. Most significantly—and unlike the use of a defendant’s prior convictions for impeachment purposes—other acts are admissible regardless of whether the defendant testifies in his own defense.¹⁹

offenders “learn to identify those factors that create their anger and role-play ways to competently use self-control techniques”).

15. *Michelson*, 335 U.S. at 475–76 (emphasis added). Worse yet, “[i]f the criminal justice system convicts people based on who they are, not what they have done, *we are all at risk* . . . [and w]e increase the risk of reinforcing errors based on fallacious stereotypical judgment.” Eads et al., *supra* note 14, at 176 (emphasis added).

16. See Eads et al., *supra* note 14, at 177 (“[S]ome states have recognized for decades the lustful disposition exception to the general ban on character evidence in cases involving sexual assault.”). In Wisconsin, for example, the prosecutor may use a defendant’s prior, first-degree sexual assault conviction “as evidence of the person’s *character* in order to show that the person *acted in conformity therewith*.” WIS. STAT. § 904.04(2)(b)(2) (2017) (emphasis added).

17. See Ted Sampsell-Jones, Spreigl *Evidence: Still Searching for a Principled Rule*, 35 WM. MITCHELL L. REV. 1368, 1372 (2009) (arguing that Minnesota’s rule on other-acts evidence “only pretends to enforce some ill-defined prohibition on character evidence”).

18. *United States v. Meyers*, 550 F.2d 1036, 1044 (5th Cir. 1977).

19. For a discussion of the overwhelming variety of other-acts evidence, see DANIEL BLINKA, EVIDENCE—WISCONSIN PRACTICE SERIES § 404.602 (4th ed. 2017). One of the most common types of other-acts evidence is a defendant’s prior conviction for a completely unrelated

The rules governing the admissibility of other-acts evidence are highly nuanced, vary significantly from state to state, and are inconsistently applied even within states.²⁰ In most jurisdictions, admissibility is generally governed by a rule very similar to Federal Rule of Evidence 404(b), Crimes, Wrongs, or Other Acts, which states:

(1) Prohibited Uses. Evidence of a crime, wrong, or other 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.

(2) Permitted Uses This evidence may be admissible for another purpose, such as proving motive, opportunity, intent, preparation, plan, knowledge, identity, absence of mistake, or lack of accident.²¹

These enumerated, permissible purposes are not exclusive. In fact, except for the prohibition on using such evidence to prove character, there is virtually no limit on how the creative prosecutor can use a defendant's other acts at trial.²²

Returning to our earlier battery example where the defendant claims self-defense, the prosecutor could simply offer the previous battery conviction not to prove the defendant's character trait for being hot-tempered, but rather to demonstrate his identity as the perpetrator; his

but similarly-named crime. *See, e.g.*, WIS. STAT. § 904.04(2)(b)(1) (Other-acts evidence "is admissible without regard to whether the victim of the crime that is the subject of the proceeding is the same as the victim of the similar act.").

20. *See, e.g.*, Sampsell-Jones, *supra* note 17, at 1372 ("The current [other-acts] doctrine in Minnesota is a Potemkin village."); *see also* Robert Cameron, *The Modified Just Rule: A New Standard for the Admissibility of Evidence of Other Crimes, Wrongs or Acts Under Rules 403 and 404(b) of the Montana Rules of Evidence*, 53 MONT. L. REV. 133, 133 (1992) ("In Montana other crimes evidence has been subject to widely varying standards of admissibility.").

21. FED. R. EVID. 404(b). A defendant's crimes, wrongs, or other acts may also be admissible under other rules of evidence, including FED. R. EVID. 413, 414, 608(b), 609. These rules may admit the same evidence but in varying level of detail, for different purposes, or through different methods of proof. *See infra* note 36 and accompanying text for further discussion of using a defendant's prior conviction as impeachment evidence under FED. R. EVID. 609 and comparable state statutes.

22. States that follow the so-called inclusionary approach allow the use of other acts for *any* purpose other than character. *See* State v. Hunt, 666 N.W.2d 771, 787 (Wis. 2003) (allowing other acts to show "context" and "the victim's state of mind," as well as "to corroborate information provided to the police" and "to establish the credibility of victims and witnesses"); David F. Guldenschuh, *Federal Rules of Evidence – Rule 404(b) Limits the Admission of Other Crimes Evidence, Under an Inclusionary Approach, to Cases Where It Is Relevant to an Issue in Dispute*, 55 NOTRE DAME L. REV. 574, 574 (1980).

intent to cause bodily harm; and the lack of accident in causing bodily harm. Further, the prosecutor can also use the previous battery conviction to demonstrate that the defendant did not act in self-defense.

While early courts cautioned that other-acts evidence should be used “sparingly and only when reasonably necessary,”²³ the floodgates have since been opened wide. The courts’ original position on other-acts evidence “has been remolded and chiseled down in recent years to the point that this once well-settled exclusion now serves as more of an exception rather than the rule.”²⁴ Today, once the prosecutor offers the other acts ostensibly for a permissible purpose, the “concrete result is that propensity evidence is regularly allowed in the guise of 404(b) . . . evidence.”²⁵ More to the point, other-acts evidence is simply “character evidence in disguise.”²⁶

In reality, a defendant’s other acts are usually far more probative of his character than any of the delineated statutory purposes. For example, in the hypothetical battery case, would the defendant’s unrelated, prior battery conviction really prove his identity, intent, lack of accident, or absence of self-defense in the current battery case? Usually it does not. Typically, the other act involved different facts and circumstances, was committed against a different person, and occurred months or even years earlier—all of which dramatically limit its relevance to the charged crime.

In fact, courts “routinely admit [other-acts] evidence whose relevance depends primarily on propensity so long as it ultimately goes to prove one of the listed ‘other purposes’ in 404(b).”²⁷ It does not matter that the other-acts evidence goes to the defendant’s character; as long as the prosecutor is able to articulate one of the permissible purposes in addition to character, the court will likely admit the evidence.²⁸

23. *State v. Murphy*, 524 N.W.2d 924, 928 (Wis. Ct. App. 1994) (discussing Wisconsin’s landmark case *Whitty v. State*, 149 N.W.2d 557 (Wis. 1967)).

24. Edward Pare III, *Restoring the Character Evidence Rule: Reconsidering Evidence of Crimes, Wrongs, and Other Acts in Rhode Island*, 21 ROGER WILLIAMS U. L. REV. 399, 399 (2016) (citing Paul S. Milich, *The Degrading Character Rule in American Criminal Trials*, 47 GA. L. REV. 775, 776–77 (2013)).

25. Sampsell-Jones, *supra* note 17, at 1387. While other-acts evidence is, in theory, admissible against the government’s witnesses as well, a double standard has emerged. See Jayna M. Mathieu, Note, *Reverse-Spreigl Evidence: Challenging Defendants’ Obligation to Exceed Prosecutorial Standards to Admit Evidence of Third Party Guilt*, 86 MINN. L. REV. 1033, 1034–35 (2002) (“[C]ourts tend to reach different results in *Spreigl* and reverse-*Spreigl* cases. Contrary to a *Spreigl* scenario, when defendants attempt to introduce reverse-*Spreigl* evidence, trial courts frequently exclude it.”).

26. *People v. Crawford*, 582 N.W.2d 785, 794 (Mich. 1998).

27. Sampsell-Jones, *supra* note 17, at 1371.

28. *See id.* at 1385–86.

But given the obvious prejudicial impact of other-acts evidence, coupled with the mantra that “we try cases, rather than persons,”²⁹ how can courts justify the routine admission of other-acts evidence at trial?

III. CAUTIONARY INSTRUCTIONS

In the face of the dual purpose for other-acts evidence—the impermissible character inference on the one hand and a permissible, statutory purpose on the other—courts typically deal with the character aspect of the evidence by issuing a cautionary instruction. Once a cautionary instruction “is properly given by the court, prejudice to a defendant is *presumed erased from the jury’s mind*.”³⁰

What is this cautionary instruction that has the power to instantly and permanently eliminate all traces of improper influence on the jury? It is often a short instruction, sometimes given at the time the other-acts evidence is presented, cautioning the jury that the evidence is only admissible for certain purposes such as proof of identity, intent, lack of accident, and absence of self-defense.

Cautionary instructions—often called limiting instructions or curative instructions—vary greatly across jurisdictions. The only thing the instructions have in common is that each will name the purpose or purposes for which the jury may use the evidence. Beyond that, some instructions will caution the jury that “[y]ou must not convict the defendant here because you think he is guilty of other bad conduct.”³¹ Other instructions are rather cumbersome—and probably incomprehensible to many jurors—warning that “[y]ou must not consider this act to determine the defendant’s character or character trait, or to determine that the defendant acted in conformity with the defendant’s character or character trait and therefore committed the charged offense.”³²

29. *People v. Allen*, 420 N.W.2d 499, 504 (Mich. 1988) (citing *United States v. Mitchell*, 2 U.S. (2 Dall.) 348, 357 (1795)).

30. *State v. Shillcutt*, 341 N.W.2d 716, 721 (Wis. Ct. App. 1983) (emphasis added). State court decisions throughout the country are littered with similar, but less extreme, assumptions about cautionary instructions. *See, e.g.*, *State v. Wright*, 719 N.W.2d 910, 918 (Minn. 2006) (“[A]ny potential prejudice was mitigated by the limiting instruction given to the jury.”). Further, “[u]nder the stewardship of Chief Justices Burger, Rehnquist, and Roberts, the [United States Supreme] Court has shown no hesitation in asserting that juries can and will disregard inadmissible evidence when instructed to do so.” Dan Simon, *More Problems with Criminal Trials: The Limited Effectiveness of Legal Mechanisms*, 75 L. & CONTEMP. PROBS. 167, 179 (2012).

31. MICH. MODEL CRIM. JURY INSTRUCTIONS § 4.11(3) (MICH. SUP. CT. COMM. ON MODEL CRIM. JURY INSTRUCTIONS 1993).

32. REVISED CRIM. JURY INSTRUCTIONS § 26A (ST. B. OF ARIZ. CRIM. JURY INSTRUCTION COMM. 2016).

Some instructions are incredibly brief and do little more than tell the jury the obvious: “the defendant is not on trial for a crime, wrong, or act that is not included in the indictment.”³³ Finally, other instructions are outright nonsensical and even self-contradictory. One federal court instruction tells the jury that it may consider the other acts as proof of the defendant’s intent or motive to commit the crime for which he is on trial; however, it then instructs the jury that “[y]ou may *not* consider this evidence as evidence of guilt of the crime for which the defendant is now on trial.”³⁴

Even assuming the trial judge reads a cautionary instruction that is technically accurate, internally consistent, and comprehensible to the average juror, it seems unlikely that such an instruction could erase the prejudicial impact of other-acts evidence. Jurors are not computers. They are probably not able to parse the evidence into its permissible and impermissible purposes and then disregard the impermissible purposes as if they were deleting unwanted files from a computer hard drive. That is, curative “instructions are premised on a belief in people’s ability to exert formidable control over their cognitive processing. This assumption runs contrary to the research.”³⁵

Given this gap between what the courts assume about cautionary instructions—that all prejudice will be erased from the jury’s mind—and how the human mind actually works,³⁶ we decided to put the other-acts

33. FLA. STANDARD JURY INSTRUCTIONS IN CRIM. CASES § 2.4 (FLA. B. 2016).

34. MODEL CRIM. JURY INSTRUCTIONS NINTH CIR. (NINTH CIR. JURY INSTRUCTIONS COMM. 2010) § 2.10 (emphasis added).

35. Simon, *supra* note 30, at 178 (discussing numerous studies of the effectiveness of instructions to ignore evidence or to use evidence only for limited purposes).

36. Much of the research examines cautionary instructions for the impeachment of a defendant with his prior convictions. See Eads et al., *supra* note 14, at 198 (discussing several studies); Theodore Eisenberg & Valerie P. Hans, *Taking a Stand on Taking the Stand: The Effect of a Prior Criminal Record on the Decision to Testify and on Trial Outcomes*, 94 CORNELL L. REV. 1353, 1359 n.17 (2009) (collecting several studies); Simon, *supra* note 30, at 176–79 (discussing several studies and cases). From a psychological standpoint, impeachment with a prior conviction is similar to the admission of other-acts evidence in terms of its impact on jurors’ perception of the defendant. However, there can also be significant differences. As discussed in Part II of this Article, other-acts evidence under FED. R. EVID. 404(b) need not have resulted in a criminal conviction—or even be criminal in nature—in order to be admissible. And other-acts evidence typically involves detailed testimony by witnesses, whereas prior-conviction impeachment evidence under FED. R. EVID. 609 may consist only of the fact and name of the crime of conviction. Additionally, state-court rules on prior-conviction impeachment evidence may restrict such information even further, allowing only the fact of a conviction *without* naming the specific crime. See Anna Roberts, *Conviction by Prior Impeachment*, 96 B.U. L. REV. 1977, 1987–88 (2016). And of course, the cautionary instructions will also differ, given that impeachment evidence is used for credibility, whereas other acts are used for such purposes as identity, intent, lack of accident, and absence of self-defense, among others. FED. R. EVID. 404(b); FED. R. EVID. 608(a).

cautionary instruction to an empirical test. Our study, discussed below, was approved by Beloit College's Institutional Review Board.

IV. THE STUDY

A. *Hypothesis*

We hypothesize that other-acts cautionary instructions are not effective and that jurors who are presented with other-acts evidence will, despite receiving a cautionary instruction, use the evidence for impermissible purposes that increase the likelihood of conviction.

B. *Participants*

To test our hypothesis, we recruited 250 study participants through Amazon's Mechanical Turk, an online platform for conducting social science research.³⁷ Mechanical Turk has many advantages, including "easy access to a large, stable, and diverse subject pool, the low cost of doing experiments, and faster iteration between developing theory and executing experiments."³⁸ Further, several studies have found a high degree of similarity between the judgments and behaviors of Mechanical Turk "workers" and of participants recruited in more conventional ways, such as through university subject pools.³⁹

These 250 participants served as mock jurors by reading a case summary and rendering a verdict in a hypothetical criminal case. All participants were required to be adults and U.S. citizens. To ensure data quality, we monitored the participants and rejected those who completed the task in fewer than three minutes; we immediately replaced them with new participants in order to maintain our desired sample size. After the data were collected, we discovered one participant had not indicated whether he was a U.S. citizen; we discarded the participant's data, leaving us with a total of 249 mock jurors.

Our sample was large and diverse. Participants hailed from 42 different states. Fifty-two percent were female. Participants' ages ranged from 20 years to 73 years; the mean (average) age was 34.8 years, and the median age (50th percentile) was 31 years. The ethnic composition of the sample was also diverse: 72% non-Hispanic whites, 8% African-Americans, 6% Hispanics, 9% Asian-Americans, 4% mixed race, and 1% identifying as other. Sixty-two percent of the participants have at least a four-year college degree, while an additional 29% have completed some college. Fourteen percent reported having prior jury experience.

37. Winter Mason & Siddharth Suri, *Conducting Behavioral Research on Amazon's Mechanical Turk*, 44 BEHAV. RES. 1 (2012).

38. *Id.* at 1.

39. *Id.* at 3–4.

C. *Study Design*

All mock jurors received a written case summary of a hypothetical criminal trial. The case summary described two adults who met and interacted at a party, which resulted in an accusation of a misdemeanor fourth-degree sexual assault, i.e., the defendant's sexual touching of the alleged victim without her consent. All mock jurors received the following information: an instruction on the charged crime, including its elements; a 767-word summary of the testimony from the accuser and the defendant, who were the two witnesses; and an instruction on the state's burden of proof.

More specifically, the accuser, Emily V., testified that she met the defendant, John D., at a house party. At one point during the evening, she went into a room with John and another couple—a man and a woman. Emily had been drinking alcohol, “was slightly intoxicated,” and was sitting on a couch with John while the couple sat near them in chairs. The woman then left the room. John and the man were talking, and Emily fell asleep. When she awoke, a man was standing over her and touching her buttocks without her consent. She believed this man was John, the defendant. After she told the man to “stop it,” there was a short struggle. The man then left the room. Emily learned John's full name from someone at the party and promptly reported the incident to the police.

The defendant, John D., also testified. He admitted to drinking alcohol at the party and also to consuming other drugs earlier in the day. John acknowledged being in the room with Emily, the other man, and the other woman. After Emily fell asleep and the other couple left the room, John stood over Emily and checked on her to make sure she was okay. After she responded to him, he left the room and then left the party. He denied touching Emily's buttocks “in any way or for any purpose” and also denied struggling with her.

D. *Identity Evidence*

If a crime was committed, the identity of the perpetrator was not clear from the case summary. The accuser was intoxicated and, at one point, even fell asleep. She was also unfamiliar with the defendant; she had never met or even seen him before the night of the party and did not know his first name until she later learned it from another person. Further, there were other people at the party including the other man who was in the same room with the accuser and the defendant. Finally, the defendant denied the allegation and testified that, after checking on the accuser, he simply left the room and the party, and no assault had occurred up to that point.

Given the possibility that, if a crime occurred, someone *other than* the defendant committed it, identity was an issue in the case. With regard to

identity evidence, mock jurors were randomly assigned to one of two groups, Group A (N = 124) and Group B (N = 125). Mock jurors in Group A received a stipulation on identity, which was provided immediately after the accuser's testimony. The stipulation read as follows:

The prosecutor and the defense attorney have stipulated or agreed to the existence of certain facts, and you must accept these facts as conclusively proved. In this case, the prosecutor and defense attorney have stipulated to the following facts:⁴⁰

1. The defendant, John D. was the person on the couch with Emily V.
2. The other man that was in the room left the room just as Emily V. was falling asleep.
3. The defendant, John D. was the person standing over Emily V.
4. At no time did the other man return to the room, and at no time did any other man ever enter the room.

Group A's stipulation was designed to remove all doubt about the identity of the perpetrator. That is, if the jurors believed that a sexual assault was committed, the crime could only have been committed by the defendant. The breadth of the stipulation removes all speculation that the crime could have been committed by the other man that was in the room or by some other, unidentified man at the party. Group A also received an additional instruction, immediately before the burden of proof instruction, reminding mock jurors that the term "evidence" includes the stipulation:

Evidence – Definition and Weight

"Evidence" includes the sworn testimony of witnesses and any stipulations entered into between the parties. All witnesses in this case were sworn before testifying. You are the sole judges of the "credibility," that is, the believability,

40. This instruction is closely modeled after WIS. CRIM. JURY INSTRUCTIONS NO. 162 (UNIV. WIS. LAW SCH. 2015). The enumerated facts on which the parties agreed are, of course, specific to the facts of the case.

of the witnesses and of the weight to be given to their testimony.⁴¹

The mock jurors in Group B, however, received *different* evidence to prove identity. Group B was presented with the defendant's other act: a prior conviction for misdemeanor, fourth-degree sexual assault. This information was provided immediately after the accuser's testimony:

Testimony of Police Detective L. Hamilton

Hamilton is a police detective in a different county where the defendant, John D., used to live. Hamilton testified that he knew John, and was able to identify him in court. Hamilton testified that, about three years ago in 2013, John D. was charged and convicted of fourth degree sexual assault for sexually touching a young woman, Heather B., without her consent. John entered a plea of "no contest" and did not go to trial. Hamilton read the following portion of John's written statement from 2013 about that incident with Heather B. John had written: "I met Heather twice before, and I saw her again last night and we went back to her apartment. We had some drinks and I was giving her a backrub. Heather was pretty drunk, and so was I, but she seemed to be enjoying the backrub. I reached around and touched her chest over her clothes. After a while I reached down the back of her pants and was touching her butt. She seemed pretty into it and didn't say 'stop' so I kept doing it. I thought she was into it, but she might have been too drunk to really know what was happening. After a few minutes, though, she realized what was happening and told me to 'stop.' I didn't stop right away, because I didn't know if she really meant it. But I stopped once I was sure she was really serious. She accused me of touching her vagina, too, but I never did that. I only touched her chest and her butt."

Group B also received the same jury instruction on the definition of "evidence." However, Group B's instruction did not include the reference to the stipulation as Group B did not receive a stipulation. Instead, Group B received a cautionary instruction on the other-acts evidence. We selected the most pro-defendant cautionary instruction that we located during our research. The instruction informs the jury that it is up to them

41. This instruction is an abbreviated instruction based on a combination of two Wisconsin Criminal Jury Instructions. WIS. CRIMINAL JURY INSTRUCTIONS NOS. 103, 300 (UNIV. WIS. LAW SCH. 2015).

to decide whether the other act occurred. It also instructs them that the defendant's other act can only be used for the purpose of identity, and not for any other purpose. It also specifically warns them not to use the other-acts evidence to judge the defendant's character or to convict the defendant based on his character.⁴² The instruction, in its entirety, reads as follows:

Cautionary Instruction

Evidence has been presented regarding other conduct of the defendant for which the defendant is not on trial.

Specifically, evidence has been presented that the defendant, in 2013, committed fourth degree sexual assault against a different person, Heather B. If you find that this conduct did occur, you should consider it only on the issue of identity.

You may not consider this evidence to conclude that the defendant has a certain character or a certain character trait and that the defendant acted in conformity with that trait or character with respect to the offense charged in this case involving alleged victim Emily V.

The evidence was received on the issue of identity, that is, whether the prior conduct of the defendant is so similar to the offense charged that it tends to identify the defendant as the one who committed the offense charged.

You may consider this evidence only for the purpose(s) I have described, giving it the weight you determine it deserves. It is not to be used to conclude that the defendant is a bad person and for that reason is guilty of the offense charged.⁴³

42. Even under today's rules of evidence that have chipped away at the character-evidence prohibition, a prior misdemeanor—fourth-degree sexual assault conviction for touching a person's buttocks—would likely not be admissible to prove the defendant's character. *See, e.g.*, FED. R. EVID. 413 (permitting the use of the defendant's prior sexual assault, when it involved contact with the "genitals or anus," to prove character); *see also* WIS. STAT. § 904.04(2)(b)(2) (2017) (permitting the use of the defendant's prior, *first-degree* sexual assault conviction to prove character). Therefore, when character evidence is prohibited, as it is in the vast majority of criminal cases, the prosecutor and trial judge would have to fit the defendant's prior conviction into one of the delineated, statutory purposes such as identity.

43. WIS. CRIM. JURY INSTRUCTIONS NO. 275 (UNIV. WIS. LAW SCH. 2015).

Between the two types of identity evidence, the stipulation provided to Group A is far stronger, as it forecloses all speculation about the identity of the perpetrator and provides conclusive proof that, if a crime was committed, the defendant committed it. Such a stipulation is probably the strongest of all types of evidence that could possibly be used to prove identity.

Group B, on the other hand, received evidence of the defendant's other act, which is much weaker evidence of identity. That is, the cautionary instruction framed the issue as "whether the prior conduct of the defendant is *so similar* to the offense charged that it tends to identify the defendant as the one who committed the offense charged."⁴⁴

While the defendant's other act in the case summary was *somewhat* similar—it involved touching without consent after drinking alcohol—there were also numerous differences. For example, the other act involved a person the defendant had met on two prior occasions, whereas the charged crime involved a person the defendant had just met for the first time. The other act occurred at the victim's house where the defendant was giving her a backrub, whereas the alleged crime occurred at a house party with others present and where the accuser had fallen asleep.

Because Group A's evidence on identity (the stipulation) was conclusive, Group A should have the higher conviction rate of the two groups. That is, because the other act provided to Group B had both similarities and dissimilarities to the charged crime, it is far less reliable evidence of identity. Therefore, if the cautionary instruction is truly effective, and the jurors in Group B considered the other act "only on the issue of identity,"⁴⁵ then the conviction rate of Group A should be higher than that of Group B.

We hypothesized, however, that the cautionary instruction will *not* be effective, and that jurors will use the other acts for *impermissible* purposes such as punishing the defendant for his past wrongs; preemptively convicting him in order to prevent future bad acts; or concluding that he possesses a bad character and is therefore guilty of the charged crime. If our hypothesis is correct, then Group B's conviction rate should be higher, even though Group B received weaker evidence on the issue of identity.

E. Findings

After receiving a burden of proof instruction, each mock juror rendered a verdict of guilty or not guilty. In Group A, which received conclusive proof of identity in the form of a stipulation, 41 of 124 mock jurors returned verdicts of guilt for a group conviction rate of 33.1%. In

44. *Id.* (emphasis added).

45. *Id.*

Group B, which received evidence of the defendant's somewhat similar other act, 60 of 125 mock jurors returned verdicts of guilt for a group conviction rate of 48.0%. The conviction rate among jurors who learned of the other act (Group B) was nearly 50% *higher* than the conviction rate among jurors who did not know about the defendant's earlier conviction but instead received the conclusive stipulation on identity (Group A).

The Z test for the difference between the two proportions—33.1% and 48.0%—produced a Z score of -2.4. This result is significant at the $p < .02$ level, with an exact p -value of 0.016. The p -value measures the probability of a Type I error, or the risk of obtaining a false positive when testing a hypothesis, given the two sample sizes and the difference in conviction rates between the two groups. In plain language, we are more than 98% certain ($1 - p$) that the observed difference in conviction rates between Groups A and B is a real difference that did not occur by chance.

Additionally, after participants rendered their verdict, they reported how certain they were (on a 10-point scale) that they had made a correct decision. Mock jurors in Group B—the group that received information about the defendant's other act—were more certain (mean score of 7.0) in their verdict than were jurors in Group A (mean score of 6.4). This difference is statistically significant at the $p < .04$ level. In plain language, jurors who heard evidence of the defendant's other act felt more certain their verdict was correct, even though they were provided with less persuasive evidence on the issue of identity.

We also uncovered several subsidiary findings not directly related to the main purpose of our study: (a) women were no more likely than men to vote guilty; (b) there were no statistically significant relationships between a participant's verdict and his or her age, ethnicity, or prior jury experience; (c) across education categories, better-educated participants were less likely to vote guilty, although the trend was not pronounced; and (d) mock jurors who voted guilty, regardless of the group to which they were randomly assigned, were considerably more certain that they had made the correct decision (a mean score of 7.7 on a 10-point scale, compared to a mean score of 6.1 among participants who voted not guilty, $p < .001$).

Finally, participants also answered an attention-check question that tested their recollection of the elements of the charged crime. The question included five potential elements, only three of which were correct. The attention-check results were encouraging, as 88% of all participants correctly identified the elements of the charged crime.⁴⁶ Those who voted not guilty were correct 91% of the time, while those who voted guilty were correct 84% of the time. This difference is

46. Our standard for a correct answer was high; a mock juror who identified the correct elements of the charged crime, but also an incorrect element, was classified as "incorrect."

marginally significant ($Z = -1.7, p = .09$); it suggests that those mock jurors who paid closer attention to the legal elements of the charged crime (fourth-degree sexual assault) were less likely to convict.

V. DISCUSSION: THE IMPACT OF OTHER ACTS

After conceding that jurors are likely to use other-acts evidence for impermissible purposes, courts typically permit such evidence and address its dangers by issuing a cautionary instruction.⁴⁷ The courts simply assume that this cautionary instruction will cure all unfair prejudice to the defendant. Our findings, however, strongly support our hypothesis that such cautionary instructions for other-acts evidence are not effective, that jurors will consider a defendant's other acts for impermissible purposes such as character, and that such consideration will lead jurors to convict at a higher rate.

More specifically, Group A in our study convicted at a rate of 33.1%, which should have served as a ceiling on the conviction rate, as this group received a stipulation that conclusively proved the defendant's identity. There is simply no better evidence to establish the defendant's identity than a clear, all-encompassing stipulation between the parties.

However, Group B, which received less-certain evidence on identity—the defendant's somewhat similar, three-year-old other act—convicted at a rate of 48.0%. Had the cautionary instruction been effective, i.e., had the jurors considered the other act only on the issue of identity as they were instructed, Group B's conviction rate should have been no higher than Group A's. Instead, it was much higher, and the difference was highly significant ($p < .02$). Further, jurors in Group B, after learning of the defendant's prior conviction, were more confident in their verdicts ($p < .04$).

This empirical evidence debunks the common judicial assumption that a cautionary instruction on other-acts evidence will erase all prejudice from the jurors' minds. Our findings demonstrate that other-acts evidence can lead jurors to convict a defendant not for what he has done, but for who he is. Even the best, most comprehensive other-acts cautionary instruction is about as effective as “throw[ing] a skunk into the jury box” and “instruct[ing] the jury not to smell it.”⁴⁸

Now that we have empirically confirmed our hypothesis and demonstrated the ineffectiveness of cautionary instructions for other-acts evidence, the next question is: What can defense lawyers do to protect a

47. See *State v. Payano*, 768 N.W.2d 832, 862 (Wis. 2009) (“[P]recedent suggests that cautionary jury instructions can go a long way in limiting the unfair prejudice that may result from the admission of other acts evidence.”).

48. *Dunn v. United States*, 307 F.2d 883, 886 (5th Cir. 1962).

defendant's due process and other constitutional rights when the state attempts to introduce other-acts evidence at trial?

VI. PRETRIAL STRATEGIES FOR THE DEFENSE

Litigating other-acts evidence requires an in-depth knowledge of the particular state's applicable rules and case law.⁴⁹ However, the empirical findings, case law, and secondary sources discussed in this Article point to three, interrelated strategies for defense counsel's consideration.

A. *Relevance*

Defense counsel should consider arguing that the defendant's other acts are not relevant and should therefore be excluded. Relevance can be broken down into at least two parts.

One aspect of relevance is whether the *permissible purpose*—the purpose for which the other-acts evidence is being offered—is in dispute. For example, if the state is prosecuting a defendant for domestic violence against his or her spouse, then identity will not be a contested issue at trial. Therefore, the defendant's prior battery conviction should not be admissible for that purpose. Stated more broadly, “[i]f the state offers other crimes evidence as relevant to specific elements of the crime charged, and those elements are not at issue, the other crimes evidence is inadmissible.”⁵⁰

However, not all jurisdictions follow this rule. Instead, some simply ignore this aspect of the relevancy analysis and permit the state to use other-acts evidence even for issues that are not contested.⁵¹ To make matters more complicated, the laws vary not only across states but sometimes within states.⁵² For example, one court conceded that “[t]he conflict between our decisions . . . poses a question of the precedent to be followed.”⁵³

49. See, e.g., Ruth Miller, *Other Crimes Evidence: Relevance Reexamined*, 16 J. MARSHALL L. REV. 371, 385–88 (1983) (discussing the various standards used in Illinois); Pare III, *supra* note 24, at 399–402 (arguing for the “clear and convincing” standard instead of the “preponderance of the evidence” standard in Rhode Island). Not only does the law in each state and federal jurisdiction vary, but this Article is not intended to be comprehensive with regard to litigating other-acts motions. There are several issues not even addressed in this Article, including the notice requirement for the use of other-acts evidence and the burden of proof required to demonstrate that the other acts actually occurred.

50. *State v. Bedker*, 440 N.W.2d 802, 805 (Wis. Ct. App. 1989).

51. See *Sampsell-Jones*, *supra* note 17, at 1400.

52. See *id.* at 1372 (“[T]he history of the rule in Minnesota and elsewhere is a history of substantial confusion. . . . [T]he rule itself has been enforced inconsistently throughout its history.”).

53. *State v. Clark*, 507 N.W.2d 172, 175 (Wis. Ct. App. 1993); see also *Sampsell-Jones*, *supra* note 17, at 1372 (“[T]he history of the rule in Minnesota and elsewhere is a history of

The second aspect of relevancy is whether the *other acts* are relevant to establish the stated, permissible purpose. For example, if the state prosecutes a defendant for a masked robbery where the identity of the perpetrator is truly in dispute, the question then becomes whether other-acts evidence (such as the defendant's prior robbery conviction) is relevant in establishing the defendant as the perpetrator. This analysis typically hinges on the similarity between the other act and the charged crime: the greater the similarity, the more likely the other act is to be relevant and, therefore, admissible.⁵⁴ One problem in determining whether two things are similar, however, is the vagueness of the inquiry:

[I]n many cases, there is no way to determine which factors cut which way. If a man [commits a crime] in Duluth and another in Minneapolis, are the [crimes] geographically similar because they both took place in the same state, or are they geographically different because they took place 150 miles apart? Asking that question is roughly equivalent to asking whether I am similar to a chimp.⁵⁵

To apply this relevance analysis to the hypothetical sexual assault case used in our study, there were several significant differences between the other act and the charged crime. For example, in the other act, the defendant had met the accuser on prior occasions; in the charged crime, the defendant met the accuser only a short time before the alleged assault. In the other act, the defendant and the accuser were alone at the accuser's home while the defendant gave her a backrub; in the charged crime, the defendant and the accuser were at a house party with several other people present. In the other act, the accuser was awake; in the charged crime, the accuser had fallen asleep.

However, if a court were predisposed to admit the other-acts evidence, it could find numerous similarities on which to hang its hat. For example, both the other act and the charged crime involved alcohol consumption by the defendant and the accuser. Further, each of the accusers was in an impaired state—whether due to intoxication or sleep. Finally, both the other act and the charged crime occurred in private homes, rather than in public places.

This example leads to an even larger problem: in many cases, courts do not undertake the similarity inquiry in good faith. Rather, they have decided ahead of time to admit the other acts. Then, they search for

substantial confusion. . . . [T]he rule itself has been enforced inconsistently throughout its history.”). The inconsistency of other-acts case law is not limited to Wisconsin or to this particular aspect of the analysis.

54. See, e.g., Sampsell-Jones, *supra* note 17, at 1392 (discussing Minnesota's competing “marked” and “substantial” similarity tests for the admissibility of other-acts evidence).

55. *Id.* at 1392–93.

commonalities—no matter how insignificant—to reach that predetermined conclusion.

For example, where a defendant is charged with the serious crime of “theft of a motor vehicle,” his prior bad acts of “receiving stolen property” in the form of “a television set” and other “electrical equipment” would strike most fair-minded readers as being very dissimilar to the charged crime.⁵⁶ Nonetheless, in a short opinion unencumbered by legal analysis, an appellate court upheld a trial court’s admission of the petty, other-acts evidence, holding that passively receiving a stolen television set and electrical equipment were substantially similar to stealing an automobile.⁵⁷ Why? Because the other-acts evidence and the charged crime were alleged to have occurred in the same county, and “[e]ach was an offense against property.”⁵⁸

By cherry-picking the most superficial similarity—in the above case, the other act and the charged crime were both listed in the same chapter of the criminal code—courts can justify admitting nearly any prior act offered up by the prosecutor. In fact, a review of the case law in any state will reveal that “[c]ourts have a long history of admitting evidence with a fairly low degree of similarity.”⁵⁹

Nonetheless, because there are often real and substantial differences between the other acts and the charged crime, defense counsel can usually make a strong argument that the other-acts evidence is dissimilar, and therefore not relevant.⁶⁰ Such an analysis is also a prerequisite to raising additional challenges, described below.

B. *Unfair Prejudice*

Defense counsel should consider arguing that the other acts should be excluded due to their unfair prejudice. Courts have defined unfair prejudice as the risk “that the jurors would be so influenced by the other acts evidence that they would be likely to convict the defendant because the other acts evidence showed him to be a bad man.”⁶¹ However, as

56. *State v. Schulberg*, No. C4-95-2709, Minn. App. LEXIS 1120, at *2 (Minn. Ct. App. Sept. 24, 1996).

57. *Id.* at *3–4.

58. *Id.* at *4.

59. *Sampsell-Jones*, *supra* note 17, at 1391, 1391–92 n.123 (citing numerous cases where the courts found defendants’ other acts to be sufficiently similar to the charged crimes despite their dramatic dissimilarity).

60. Each jurisdiction will have at least some published case law to support this argument. *See, e.g.*, *State v. Sullivan*, 576 N.W.2d 30, 39 (Wis. 1998) (overturning conviction because “a domestic disturbance between the defendant and his ex-wife in which they argued but there was no physical contact” was *not* sufficiently similar to the charged crime which allegedly “involved the defendant punching the complainant”).

61. *Id.* at 40.

described earlier, this unfair prejudice might be “presumed erased from the jury’s mind” once the court issues a cautionary instruction.⁶²

Attempting to exclude other-acts evidence on grounds of unfair prejudice requires at least two steps. First, because courts typically require weighing the unfair prejudice against the probative value of the evidence,⁶³ counsel should first demonstrate (as discussed in the previous Section)⁶⁴ that the other acts are not relevant. If the other-acts evidence has little or no relevance—either because the permissible purpose for which it is offered is not in dispute, or because it does not tend to prove the permissible purpose for which it is offered—then its unfair prejudice will necessarily outweigh its probative value. The evidence, therefore, should be excluded.⁶⁵

Second, because courts can easily avoid weighing unfair prejudice against the probative value by simply issuing a cautionary instruction, counsel should debunk the misconception that cautionary instructions cure unfair prejudice. This, in turn, can be accomplished from both a legal and empirical perspective. From a legal perspective, most states have at least some published cases holding that cautionary instructions were not adequate to address the risk of unfair prejudice, and defense counsel should locate and cite any factually similar cases.⁶⁶

From an empirical perspective, the findings in this study and Article—along with the findings in related studies testing the effectiveness of cautionary instructions in similar contexts⁶⁷—demonstrate that such instructions are simply not effective. Here, an application of the Fifth

62. *State v. Shillcutt*, 341 N.W.2d 716, 721 (Wis. Ct. App. 1983).

63. See Vivian M. Rodriguez, *The Admissibility of Other Crimes, Wrongs or Acts Under the Intent Provision of Federal Rule of Evidence 404(b): The Weighing of Incremental Probity and Unfair Prejudice*, 48 U. MIAMI L. REV. 451, 454 (1993).

64. See *supra* Section VI.A.

65. Even when the other-acts evidence is very similar to the charged crime, and therefore deemed to be relevant, counsel will still have an argument to exclude the evidence. The reason is that as the level of similarity increases, so does the unfair prejudice. See, e.g., Eisenberg & Hans, *supra* note 36, at 1359 (discussing an increase in conviction rates when mock jurors “learned of a defendant’s previous record for crimes similar to that charged”).

66. The most useful cases for this purpose are those where courts ultimately deemed the cautionary instruction inadequate because the other-acts evidence was highly inflammatory. In other cases, however, cautionary instructions can be deemed inadequate simply because the judge failed to competently draft the instruction. E.g., *Old Chief v. United States*, 519 U.S. 172, 176 n.2 (1997) (criticizing the trial court’s instruction to the jury to use the defendant’s prior conviction to decide his “believability as a witness” when, in fact, the defendant never testified); *Sullivan*, 576 N.W.2d at 40 (“[I]n this case the cautionary instruction to the jury about the other acts evidence was too broad, and its cautionary effect was significantly diminished.”).

67. See *supra* note 36 and accompanying text.

Circuit's skunk analogy⁶⁸ is useful. Our findings demonstrate that even a well-drafted cautionary instruction is about as effective as throwing a skunk into the jury box and then instructing the jury to consider the smell only on the issue of whether the skunk expelled sulfur-laden chemicals, but not to conclude that the skunk is a smelly animal. To assume the jury would—or even could—follow such an instruction is “unmitigated fiction.”⁶⁹

C. *The Stipulation*

A third potential defense strategy is to stipulate to the element of the crime for which the other-acts evidence is being offered. In our study's hypothetical sexual assault case, the defendant stipulated to the facts necessary to establish his identity. If a crime was in fact committed, the defendant conceded that he would have been the one who committed it. Such an approach would provide the state with conclusive proof on an element of the crime,⁷⁰ thus rendering the other acts inadmissible. In some respects, this strategy circles back to the first two issues: the stipulation renders the other-acts evidence completely irrelevant; and, without any relevance, the probative value of the other acts would necessarily be substantially outweighed by their unfair prejudice.⁷¹

However, just as with the relevancy analysis, not all courts permit the defendant to use this strategy. Instead, some allow the prosecutor to reject the stipulation in order to “prove his case his own way.”⁷² Stated more cynically, some states allow the prosecutor to harness the unfair prejudice of the other-acts evidence to win a conviction. And, to make matters more complicated, the law not only varies across states but, once again, can also vary within states.⁷³

68. *Dunn v. United States*, 307 F.2d 883, 886 (5th Cir. 1962) (analogizing the cautionary instruction to telling the jurors to ignore the smell of a skunk just thrown into the jury box).

69. *Krulewich v. United States*, 336 U.S. 440, 453 (1949) (Jackson, J., concurring) (discussing the ineffectiveness of cautionary instructions).

70. For this reason, stipulations should not be offered casually, and only after consulting with the defendant and complying with all of the state's procedural safeguards.

71. See *Rodriguez*, *supra* note 63, at 456 (“The scaling of probative value and potential unfair prejudice . . . should take into consideration the defendant's offer to remove the issue . . . from the case.”).

72. *Old Chief v. United States*, 519 U.S. 172, 177 (1997). Conversely, the Second Circuit has held that the defendant's stipulation may prevent the government's use of such evidence. See *United States v. Manafzadeh*, 592 F.2d 81, 87, 88 (2d Cir. 1979). This, however, is the minority view. See, e.g., *United States v. Chaimson*, 760 F.2d 798, 805 (7th Cir. 1985) (“This court has previously rejected the Second Circuit's reasoning in *Manafzadeh* because, in effect, it allows a defendant to remove intent as an element of the crime charged.”).

73. Here, once again, Wisconsin serves as an example of the legal chaos, as two lines of authority emerged: one where the defendant has the right to stipulate to an element of the crime and avoid the state's other-acts evidence and another where he does not. See *State v. Veach*, 648

In any case, the analysis for stipulations begins with *Old Chief v. United States*.⁷⁴ Pursuant to *Old Chief*, the defendant has the right, in some circumstances, to force the trial judge and the prosecutor to accept a stipulation.⁷⁵ Further, defense counsel could even be found ineffective for failing to offer one.⁷⁶ But this per se rule may be limited to situations where the evidence is being offered to prove a defendant's legal status—for example, his preexisting status as a convicted felon when charged with being a felon in possession of a firearm⁷⁷—which is typically distinct from the classic uses of other-acts evidence.⁷⁸

When the prosecutor offers other acts for purposes other than proving the defendant's legal status, the trial judge or the prosecutor might be free to reject a defendant's stipulation. For example, in *State v. Veach*,⁷⁹ the court cited *Old Chief* and held that the trial judge properly rejected the stipulation because, rather than being used to prove the defendant's legal status, the other-acts evidence was being offered to demonstrate his "intent or motive" or absence of "mistake."⁸⁰ The court seemed to justify this ruling on two different grounds, neither of which is persuasive.

First, the *Veach* court stated that a defendant "may not stipulate or admit his way out of the full evidentiary force of the case as the Government chooses to present it."⁸¹ But if this reasoning was valid, courts would never bother issuing a curative instruction, the purpose of which *is* to limit the full evidentiary force of other-acts evidence. In other words, the risk that accompanies other acts is that they could be used as character evidence. But the problem is not that "character is irrelevant; on the contrary," a defendant's character *is* relevant.⁸² The reason such evidence is excluded is the fundamental principle that "we try cases,

N.W.2d 447, 474 (Wis. 2002) (Crooks, J., concurring) ("The majority implies that it is overruling *Wallerman* and *DeKeyser* . . . but does not do so explicitly. In fact . . . the majority suggests that it is only modifying those cases by stating, 'We do not mean to imply that *Wallerman* stipulations are per se invalid').

74. 519 U.S. at 174 (1997).

75. *See id.* (holding that, in limited circumstances, it is an abuse of discretion for a trial court to reject a defendant's stipulation).

76. *See, e.g., State v. Rivera*, No. 2014AP2302-CR, 2015 (Wis. Ct. App. Sept. 17, 2015) (per curiam) (holding that trial counsel was ineffective for failing to stipulate that the defendant had been charged with a felony at the time he allegedly committed the act that formed the basis for his bail jumping charge).

77. *See Old Chief*, 519 U.S. at 174.

78. *See Rodriguez, supra* note 63, at 452.

79. 648 N.W.2d 447, 450 (Wis. 2002).

80. *Id.* at 453.

81. *Id.* at 472 (citing *Old Chief*, 519 U.S. at 186–87).

82. *Michelson v. United States*, 335 U.S. 469, 475–76 (1948).

rather than persons.”⁸³ And because a jury should only convict a defendant “for what he did, not for who he is,”⁸⁴ the defendant therefore *should* have the right to evade the full evidentiary force of the other-acts evidence.⁸⁵

And this point leads us back to our empirical findings: cautionary instructions for other-acts evidence are not effective. Jurors will consider a defendant’s other acts for impermissible purposes such as character, and such consideration does lead jurors to convict at a much higher rate. Yet, this problem would be solved if the trial judge and the prosecutor were required to accept a defendant’s offer to stipulate to the element for which the other acts are ostensibly being used.

Second, the *Veach* court also stated that, unlike the stipulation in *Old Chief*, “the stipulation proposed by the defendant is simply inadequate to inform the jury of what is agreed to and what is in dispute, and to remove the issues from the case.”⁸⁶ This can be a legitimate point. A defendant’s stipulation should be broad enough to completely resolve the issue. For example, the stipulation used in our study removed all doubt about identity, and the jury could only conclude that, if a crime was in fact committed, the defendant committed it.

But drafting a stipulation broadly enough to satisfy a court may require even more creativity. In *Veach*, the defendant was charged with sexually touching a child and offered to stipulate that, if the touching occurred, it “was *intentional* and for the purpose of sexual gratification.”⁸⁷ But the court concluded that such a stipulation was defective, as it “would not properly inform the jury that *accident* or *mistake*, two issues which the facts of this case obviously touched upon, were subject to the stipulation.”⁸⁸

The court’s decision can only be described as disingenuous. If a defendant stipulates that an act “was intentional and for the purpose of sexual gratification,” the act is, by definition, not an “accident or mistake.” As the three dissenting judges stated, instead of being “honest

83. *People v. Allen*, 420 N.W.2d 499, 504 (Mich. 1988) (citing *United States v. Mitchell*, 2 U.S. (2 Dall.) 348, 357 (1795)).

84. *United States v. Meyers*, 550 F.2d 1036, 1044 (5th Cir. 1977).

85. Some “other acts” do have legitimate evidentiary force that should not be diluted by a stipulation. However, such evidence does not fall into the category of other-acts evidence. Rather, because such acts are “so inextricably intertwined with, or intricately related to, the charged conduct that they help the fact-finder form a more complete picture of the crime,” they are considered an entirely separate class of evidence and are governed by different rules of admissibility. *United States v. Conner*, 583 F.3d 1011, 1018 (7th Cir. 2009); *see also* Schuster, *supra* note 8, at 961–70 (discussing the “inextricably intertwined” doctrine).

86. *Veach*, 648 N.W.2d at 473.

87. *Id.* (emphasis added).

88. *Id.* (emphasis added).

and forthright . . . the majority engages in legal gymnastics to justify the admission of propensity evidence in contravention of the statute.”⁸⁹

The dissenters’ harsh words ring true and provide the only plausible explanation for the majority’s pseudo-reasoning. They also provide a cautionary tale for defense counsel: If the defense decides to offer a stipulation on an element of the crime, the stipulation should be both broad and redundant.⁹⁰

VII. POTENTIAL CRITICISMS OF THE STUDY

Our findings provide strong evidence that other-acts cautionary instructions are not effective, that jurors will use a defendant’s other acts for impermissible purposes, and that jurors exposed to other-acts evidence will convict at a much higher rate. Our findings also make a strong case for the expansion of *Old Chief*’s holding. That is, our justice system should permit a criminal defendant to stipulate not only to his legal status but also to other elements of the charged crime in order to avoid the unfairly prejudicial impact of other-acts evidence.

However, critics may argue that our study has several limitations: our use of the case summary method, our use of a single criminal charge and fact pattern, the lack of deliberation by our mock jurors, our inability to observe the participants’ level of attention, and our failure to screen the participants for bias. We discuss these potential criticisms below.

A. Case Study Method

Our study employed a written case summary method, in which the jury was provided with the elements of the charged crime, a summary of the witnesses’ testimony, and jury instructions. This is similar to the method that has been used in many peer-reviewed studies, including studies that examined the impact of a defendant’s personal characteristics on jury decision making.⁹¹ However, some social scientists have

89. *Id.* at 475 (Abrahamson, C.J., Bablitch & Bradley, JJ., dissenting).

90. One creative way of attempting to ensure that a stipulation is adequately broad is to remove the element of the crime from the jury instruction itself. To draw an analogy, when trial judges fail to instruct jurors on all of the elements of the charged crime, courts have held that such omissions can be harmless error. *See, e.g.,* *People v. Flood*, 957 P.2d 869, 872 (Cal. 1998) (stating that “removing [an] element of the crime from the jury’s consideration” does not invalidate the conviction). Therefore, eliminating elements of the crime from the jury instruction for purposes of a stipulation, when done on the defendant’s motion, should be legally permissible. Ensuring the stipulation is adequately redundant, however, may require an even greater level of imagination.

91. Wayne Weiten & Shari S. Diamond, *A Critical Review of the Jury Simulation Paradigm*, 3 *LAW & HUM. BEHAV.* 71, 77 (1979); *see also* Lawrence T. White, *Juror Decision Making in the Capital Penalty Trial: An Analysis of Crimes and Defense Strategies*, 11 *LAW & HUM. BEHAV.* 113, 129–30 (1987).

expressed concerns about the ecological validity of these studies and have called for more realistic trial simulations.⁹²

Conversely, other social scientists have noted the prohibitive costs of more realistic trial simulations⁹³ or have failed to observe differences in the reactions of mock jurors to abbreviated or more elaborate case summaries.⁹⁴ Therefore, “[e]ven highly artificial simulations are not *inherently* distorting and may actually inform us on relationships of real significance for law and human behavior.”⁹⁵ Further, the more realistic trial simulation methods actually “provide a myriad of additional legally relevant and *irrelevant* bases on which to make a decision,” including, for example, a defendant’s or witness’s race and ethnicity.⁹⁶

Seen in this light, the simplicity of the case summary method may actually be its strength. First, researchers who use the case summary method can eliminate extrajudicial factors, including race and ethnicity, that may have an impact on jurors’ decision-making processes.⁹⁷ Second, the more abbreviated case summary method compresses events in time, thereby reducing the pernicious effect of forgetting, which can also affect jurors’ decision-making processes.⁹⁸ And third, the case summary method allows researchers to test the impact of a specific component of a trial that may get lost in the clutter of a more complex trial simulation.⁹⁹ For these reasons, “more abbreviated experimental stimulus materials,” such as the case summary method, “can play an important role in addressing some questions about jury behavior.”¹⁰⁰

92. Joel D. Lieberman & Bruce D. Sales, *What Social Science Teaches Us About the Jury Instruction Process*, 3 PSYCHOL. PUB. POL. & LAW 589, 592 (1997) (“A better methodology is to provide a videotaped trial to participants. The videotaped format provides a highly engaging simulation, and is much more representative of an actual trial. Consequently, greater faith can be placed in studies using this methodology than experiments using transcripts or case summaries.”).

93. Shari S. Diamond, *Illuminations and Shadows from Jury Simulations*, 21 LAW & HUM. BEHAV. 561, 566 (1997).

94. Geoffrey P. Kramer & Norbert L. Kerr, *Laboratory Simulation and Bias in the Study of Juror Behavior: A Methodological Note*, 13 LAW & HUM. BEHAV. 89, 89 (1989) (“Results provided no support for the contention that treatment effects act differently as a function of the length of the stimulus trial in which they are embedded. Rather, it is suggested that treatments used in simplified jury simulations may often show similar effects when examined in more realistic, complex settings if the treatments are comparable.”).

95. *Id.*

96. Diamond, *supra* note 93, at 564 (emphasis added).

97. *Id.*

98. *Id.* at 563–64.

99. *Id.* at 564.

100. *Id.* (arguing that while certain studies, such as those that test “the credibility of various types of expert testimony” demand “a fairly elaborate simulation,” other studies can be accomplished using “a less extensive trial stimulus”).

B. *Single Fact Pattern*

All empirical studies are flawed in the sense that methodological decisions designed to solve one problem often exacerbate another. For example, controlled experiments that use random assignment, such as ours, solve the problem of causal ambiguity, i.e., determining what produced the effect. However, the desire to control extraneous variables may constrain the researcher's ability to generalize the study's results beyond the specific conditions tested.

We used a single fact pattern that we held constant across test conditions. Consequently, we cannot assume an identical outcome for different fact patterns. We therefore encourage future researchers to replicate our study but to use different test materials when doing so, including, for example, a different charged crime, fact pattern, other-acts evidence, and permissible purpose for the other-acts evidence.

C. *Lack of Deliberations*

Our study tested the impact of the cautionary instruction on individual mock jurors' verdicts. These mock jurors did not deliberate as a group before reaching their decisions. In this sense, our study differed from an actual jury trial and some other jury simulation studies.

Some studies show that "deliberations sometimes do influence outcomes" including, for example, a study in which juror deliberations reduced individual juror biases and made them more likely to follow the judge's instructions.¹⁰¹ However, the evidence on the impact of deliberations is, at best, mixed.¹⁰²

For example, several studies have tested the impact of deliberations on the physical attractiveness bias, i.e., the tendency for jurors to perceive and treat attractive defendants more favorably than plain-looking defendants.¹⁰³ One study found that deliberation *mitigated* the physical attractiveness bias.¹⁰⁴ A later study, however, found that deliberation *exacerbated* the bias.¹⁰⁵ Most surprising of all, a third study found a

101. *Id.* at 565.

102. *Id.*; see also Lieberman & Sales, *supra* note 92, at 635 ("On the basis of these contradictory findings, we cannot assume that deliberation will eliminate the problem of incomprehensible instructions.").

103. Marc W. Patry, *Attractive but Guilty: Deliberation and the Physical Attractiveness Bias*, 102 PSYCHOL. REP. 727, 727 (2008).

104. Richard R. Izzett & W. Leginski, *Group Discussion and the Influence of Defendant Characteristics in a Simulated Jury Setting*, 93 J. SOC. PSYCHOL. 271, 271 (1974).

105. Robert J. MacCoun, *The Emergence of Extralegal Bias*, 17 CRIM. JUST. & BEHAV. 303, 303 (1990).

reversal of the expected effect: Mock jurors who deliberated were biased against the attractive defendant.¹⁰⁶

Numerous published studies involving mock jurors do not include deliberations.¹⁰⁷ Further, in the end, requiring mock jurors to deliberate before rendering a verdict is unlikely to change the observed pattern of verdicts across conditions. Rather, “[t]he prevailing view . . . is that deliberations play a minor role in determining jury verdicts because the predeliberation majority generally prevails in the end.”¹⁰⁸

D. *Participant Attention Level*

When using Mechanical Turk, as opposed to collecting data in a laboratory setting, it is not possible to directly observe the study participants’ level of attention. However, before results were known, we were able to reject participants who spent fewer than three minutes on the task. We also tested our mock jurors’ attention level through the use of a post-verdict attention-check question. As indicated above, these results were encouraging. Mock jurors answered the question correctly 88% of the time, thus demonstrating that they devoted adequate attention to the case study materials before rendering their verdicts.

Furthermore, the issue of inattentive jurors is a problem that exists with real-life jurors as well. Far worse than merely being inattentive, sleeping jurors are often tolerated as long as the trial judge concludes that the jurors were not sleeping too long, or that the evidence they missed was not important enough, to justify a new trial.¹⁰⁹

E. *Participant Bias*

In theory, biased jurors are excused from jury duty and do not participate in actual trials. In our Mechanical Turk study, we were not able to screen participants in advance for potential bias. However, four things mitigate this potential problem. First, many biases—for example, racial bias—would not have been factors in our case. Our hypothetical defendants’ (and even the witnesses’) race was not provided.

106. Patry, *supra* note 103, at 731.

107. *See, e.g.*, Bette L. Bottoms et al., *Gender Differences in Jurors’ Perceptions of Infanticide Involving Disabled and Non-Disabled Infant Victims*, 35 CHILD ABUSE & NEGLECT 127, 131 (2011); Joanna D. Pozzulo et al., *The Effects of Victim Gender, Defendant Gender, and Defendant Age on Juror Decision Making*, 37 CRIM. JUST. & BEHAV. 47, 54 (2010); Cynthia J. Najdowski et al., *Jurors’ Perceptions of Juvenile Defendants: The Influence of Intellectual Disability, Abuse History, and Confession Evidence*, 27 BEHAV. SCI. & LAW 401, 406, 415 (2009).

108. Diamond, *supra* note 93, at 564.

109. *See State v. Chestnut*, 643 S.W.2d 343, 346, 347 (Tenn. Ct. App. 1982) (upholding the denial of defendant’s motion for new trial despite undisputed evidence of three jurors sleeping through evidentiary portions of trial).

Second, in actual courtrooms across the country, biased jurors find their way onto juries. In fact, to exclude a *subjectively* biased juror, all of the following must happen: the would-be juror must be aware of his or her bias; the judge or the attorneys must devise questions to expose that particular bias; and the would-be juror must actually admit his or her bias to a roomful of fellow citizens. For all of these things to happen is a rare occurrence indeed.

Third, even when a juror is, by all accounts, *objectively* biased, he or she may still be permitted to serve on the jury. Perhaps the most egregious example occurred when a court permitted the prosecutor's own employee to serve on the defendant's jury, finding that the employee–employer relationship between the juror and the prosecutor was not sufficient to justify the juror's removal.¹¹⁰

Finally, there is a fourth mitigating factor: the random assignment of the study participants to test conditions. The virtue of random assignment is that, when used with large numbers of study participants, it produces groups that are statistically equivalent to each other in all respects. Each group has roughly the same number of mock jurors, the same number of men and women, the same number of well-educated and poorly educated persons, and the same number of biased and unbiased individuals.

When test groups are statistically equivalent at the outset; receive different evidence on identity; and then convict at different rates, we can be quite certain that the different conviction rates were produced by the different evidence and not by personal characteristics of the mock jurors in a particular group. In plain language, random assignment creates a level playing field where the effects of bias are distributed equally across the test conditions. Therefore, we can attribute the result—a difference in conviction rates—only to the variable that was manipulated.

CONCLUSION

While evidence of a defendant's bad character is generally not admissible at trial, courts will still permit the prosecutor to introduce evidence of the defendant's prior bad acts. In some jurisdictions, the

110. See *State v. Smith*, 716 N.W.2d 482, 493 (Wis. 2006). The dissent, however, offered a far more rational view:

An objectively reasonable person in the place of the challenged prospective juror would not ordinarily be able to separate his or her economic and loyalty interests from the determinations he or she would be required to make as juror. An employee of a district attorney's office should therefore be struck as a juror for cause when that office is prosecuting a case.

Id. at 495 (Abrahamson, C.J., dissenting). Therefore, this problem—the potential for participant bias—again mirrors the problems encountered with real-life jurors.

prosecutor's only substantive hurdle is that the other-acts evidence must also be relevant for some permissible purpose—such as the defendant's identity, intent, or absence of accident—in addition to demonstrating bad character.¹¹¹

In order to protect the defendant against the impermissible character inference, courts will issue a cautionary instruction telling the jury to consider the other acts only for their legally proper purpose. Courts then blindly assume that such cautionary instructions will erase all prejudice from the jurors' minds.¹¹²

Given that this judicial assumption does not square with the psychological research on the way the human mind functions, we empirically tested this claim. We recruited 249 mock jurors, all of whom read the same case summary of a hypothetical criminal trial. Mock jurors were then randomly assigned to one of two groups, each of which received different evidence on the issue of identity.¹¹³

Mock jurors in Group A received *conclusive* evidence on identity in the form of a stipulation. One-third of these mock jurors convicted the defendant. Mock jurors in Group B received *less certain* evidence on identity in the form of the defendant's somewhat similar other act. Group B also received a cautionary instruction. Nearly half of these mock jurors convicted the defendant, even though they were instructed that the evidence could be used only for the purpose of identity and not to determine the defendant's character. This nearly 50% increase in conviction rates was highly significant ($p < .02$). Jurors in Group B (who received the other-acts evidence) were also more certain in their verdicts ($p < .04$).¹¹⁴

Our findings are strong empirical evidence that other-acts evidence increases the likelihood of conviction and that cautionary instructions for other acts are not effective. Given this, defense counsel should consider offering a stipulation to the element of the crime for which the other-acts evidence is ostensibly being offered. Further, defense counsel should consider using the empirical findings discussed in this Article—with or without the use of a stipulation—to demonstrate that the unfair prejudice associated with other-acts evidence substantially outweighs any probative value and, therefore, the other acts should be excluded from the defendant's trial.¹¹⁵

111. See *supra* Part II.

112. See *supra* Part III.

113. See *supra* Sections IV.A–IV.D.

114. See *supra* Section IV.E.

115. See *supra* Part VI.

