Character at the Crossroads

Roger C. Park

UC Hastings College of the Law, parkr@uchastings.edu

Follow this and additional works at: http://repository.uchastings.edu/faculty_scholarship

Part of the Evidence Commons

Recommended Citation

Available at: http://repository.uchastings.edu/faculty_scholarship/594
Author: Roger C. Park
Source: Hastings Law Journal
Citation: 49 Hastings L.J. 717 (1998).
Title: Character at the Crossroads

Originally published in HASTINGS LAW JOURNAL. This article is reprinted with permission from HASTINGS LAW JOURNAL and University of California, Hastings College of the Law.
Character at the Crossroads

by

ROGER C. PARK*

Table of Contents

I. The Basic Contours of the Character Evidence Ban .................. 718

II. Rationales for Exclusion of Character Evidence .................. 720
   A. Preventing Mistake and Misuse by the Trier of Fact .......... 720
      (1) Inferential Error Prejudice ........................................ 720
         a. The Probative Value of Character Evidence .............. 721
            (i) Two Illustrations of Probative Value .......... 725
            (ii) Character as Bunk? The Challenge of
                 Situationism ............................................. 728
         b. Flaws in Human Reasoning That Might Cause Misuse of Character Evidence .................. 738
         c. Problems Caused by the Trial Setting ................. 741
         d. The Problematic Nature of A Priori Regulation of
            Inferential Error Prejudice ................................ 743
               (i) Difficulty of Setting Rules Ahead of Time ... 743
               (ii) The Questionable Assumption of Judicial
                    Superiority ............................................ 744
      (2) Nullification Prejudice ......................................... 745
   B. Cost-effectiveness .................................................. 746
   C. Preventing Adversarial Misconduct and Distortion ............ 748
      (1) Bullying, Threats, Revenge .................................. 748
      (2) Surprise, Ambush, and Attendant Hazards .............. 748
      (3) Round up the Usual Suspects ................................. 749
   D. Goals Outside the “Rationalist Tradition” ...................... 749
      (1) Departures from the Rationalist Tradition—
         as Description .................................................. 751
      (2) Departures from the Rationalist Tradition—
         as Prescription ................................................ 753

III. Particular Problems with the Current Rules ...................... 754
   A. Complexity ......................................................... 754
   B. Evasions, Misuse, and Mystification ............................ 754
   C. Inconsistent Treatment of Like Cases ............................ 755
   D. Perverse Side Effects ............................................ 756

IV. Cures and Palliatives ................................................ 756
   A. Crime-by-Crime Approach to Relaxation of the Ban ........ 756

* James Edgar Hervey Professor of Law, Hastings College of the Law, University of California. I wish to thank David Faigman, Richard Friedman, Steve Penrod, and Miguel Méndez for helpful comments, and Lynn Dean and Jim Day for able research assistance. I am indebted to Peter Tillers, whose ideas have influenced and inspired my work.
I. The Basic Contours of the Character Evidence Ban

In an often-quoted passage, McCormick wrote that "character" is "a generalized description of one's disposition, or of one's disposition in respect to a general trait, such as honesty, temperance, or peacefulness." Under this definition, the character ban prohibits evidence aimed at eliciting inferences based on attributions of broad cross-situational propensities. Clear examples include attributions that a person is:

- lawless
- law-abiding
- violent
- peaceable
- a liar
- trustworthy
- intemperate
- temperate
- a thief
- an altruist
- cruel
- kind
- lazy
- conscientious
- careless
- careful

This ban on cross-situational propensity evidence is usually car-

---

ried a step further, so that a defendant cannot be shown to have the propensity to commit a crime by use of evidence that merely shows that the defendant committed another crime that has the same statutory description. This means that a number of somewhat more specific propensities, what I will call Level Two propensities, are considered to be traits of character for purposes of the prohibition. These include the propensity to be:

- a bank robber
- a counterfeiter
- an embezzler
- a murderer
- a tax cheater
- a drug dealer
- a kidnapper
- a spouse-abuser
- a drunken driver
- a rapist

However, the legal concept of "character" does not exclude evidence of more narrowly defined propensities, such as the propensity to:

- abuse a particular spouse
- rob banks using poetic threats
- drown brides in the bath after insuring their lives
- warn dental patients of the dangers of anesthetic

In other words, if testimony leads to an inference that a person committed an act without any intermediate inference about general cross-situational propensity (character), then the character evidence ban does not apply. The ban does not apply even if the evidence shows "propensity" in the sense of showing a proclivity to repeat situation-specific conduct. Examples of propensity evidence that is not, in the eyes of the law, character evidence, include testimony about habit (e.g., plaintiff always signaled a left turn), modus oper-
HASTINGS LAW JOURNAL

andi (e.g., defendant had previously robbed banks using the same disguise), or motive (e.g., defendant had gambling debts, and therefore had a motive to embezzle).

The core of the character evidence ban is its exclusion of evidence whose value depends upon attributions of broad cross-situational propensities—inferences drawn from traits at the level of generality represented by the Level One and Level Two lists. Even for that type of evidence, there have been a few traditional exceptions—for example, the “mercy rule” allowing a criminal defendant to call character witnesses, and the exception allowing attack upon a witness’ character for truthfulness with evidence of prior convictions. Federal Rules of Evidence 413 and 414, effective in 1995, added a significant new exception allowing character reasoning in cases involving serious sex crimes.

This paper will explore the rationale for exclusion of other-crimes evidence and, in the changed environment brought by Rules 413 and 414, ask whether evidence should be admitted for character reasoning in cases other than sex-crime cases.

II. Rationales for Exclusion of Character Evidence

A. Preventing Mistake and Misuse by the Trier of Fact

(1) Inferential Error Prejudice

Character evidence raises dangers of prejudice. The prejudice feared is of two types: inferential error prejudice and nullification prejudice. The former occurs if the trier overvalues the evidence in determining whether the defendant committed the crime charged. The latter occurs if the trier decides to punish the defendant for acts other than those charged. I will start by examining the rule’s sup-

5. See FED. R. EVID. 404(a)(1); 1A WIGMORE ON EVIDENCE § 58.2 (Tillers Rev. 1983) (defense character evidence); FED. R. EVID. 609 (impeachment of character of witness with prior conviction).

6. See FED. R. EVID. 413-14. For a similar state provision, see CAL. EVID. CODE § 1108.

7. The classic statement appears in Justice Jackson’s opinion in the Michelson case: The inquiry is not rejected because character is irrelevant; on the 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. The overriding policy of excluding such evidence, despite its admitted probative value, is the practical experience that its disallowance tends to prevent confusion of issues, unfair surprise and undue prejudice. Michelson v. United States, 335 U.S. 469, 475-76 (1948) (footnotes omitted).
posed value in preventing inferential error.

a. The Probative Value of Character Evidence

Common sense (that is, unsystematic inductions from everyday experience) is the principal basis both for believing character evidence to be probative, and for fearing its prejudicial effect. We all use character evidence in ordinary life. Most of us believe that there are personality differences between those who commit outrageous crimes and other people, and that ordinary people are more restrained by morality and by fear of consequences than are criminals. The data gathered systematically by scholars of crime does not require us to abandon that common sense view. Scholarship about crime supports the view that there are individual differences between repeat offenders and the rest of the population. Recidivism data,

8. It is also valued in the marketplace. Were there no such thing as proclivity to drive poorly, an easy fortune could be made at the expense of conventional insurance companies. In fact, some automobile insurance companies reportedly even use an applicant's credit history as a way of predicting underwriting risk, on the theory that “individu-


10. See MICHAEL R. GOTTFREDSON & TRAVIS HIRSCHI, A GENERAL THEORY OF CRIME passim and 107 (1990) (“[R]esearch regularly shows that the best predictor of crime is prior criminal behavior” and that the differences between people, with respect to the likelihood they will commit criminal acts, persist over time.). See also JOHN MONAHAN, U.S. DEPT OF HEALTH AND HUMAN SERVICES, THE CLINICAL PREDICTION OF VIOLENT BEHAVIOR 71-72 (1981) (“If there is one finding that overshadows all others in the area of prediction, it is that the probability of future crime increases with each prior criminal act.”); MARK H. MOORE ET AL., DANGEROUS OFFENDERS: THE ELUSIVE TARGET OF JUSTICE 23-62 (1984) (empirical data support intuition that certain individuals are unusual in their high propensity to commit crime and high rate of offending). But cf. THOMAS GABOR, EVERYBODY DOES IT! CRIME BY THE PUBLIC (1994). Gabor, a criminologist who dislikes punitive attitudes toward crime and stereotypical thinking about criminals, concedes that people differ in their commitment to crime and in the seriousness of their transgressions, but asserts that most people violate the law at some point, and that the difference is only one of degree. See id. at 5-12. Gabor presents an eclectic collection of information suggesting that cheating, lying, stealing, and sexual imposition are widespread among ordinary people. In dealing with homicide, however, he is singularly unconvincing. He relies upon anecdotes about apparently ordinary people who committed homicide, and on data showing that many homicides (a) arise from disputes among acquaintances or family members, and/or (b) are to some degree “victim-precipitated” (the result of fights or arguments). See id. at 100-05. Information
whatever its faults and shortcomings, at a minimum shows that someone who has been convicted and imprisoned for a criminal offense is many times more likely to commit a similar offense than a person chosen at random.11

In sentencing and civil commitment proceedings, other-crimes evidence is admissible to show general propensity. There, the evidence may be used for the purpose of predicting future conduct, even though it would be prohibited in a guilt determination proceeding when offered to show past conduct. Yet propensity evidence often throws more light on the past than on the future. When past conduct of this nature does nothing to counter the common-sense view that murderers are different from ordinary people (even if they murder friends and relatives during fights).

Though prediction of future crime based on prior offenses is fraught with difficulties, there are reports of success. An index of "assaultive risk," used in Michigan to predict the danger of parole failure on the basis of four simple questions about past behavior, is reportedly able to distinguish between individuals with a 40% chance of parole failure due to assault and those with a 2% chance. A "yes" answer to three questions (whether the inmate was in prison for robbery, sexual assault or murder; whether there had been serious institutional misconduct; whether the inmate was first arrested before his 15th birthday) predicted a high (40%) risk that the inmate would be arrested and returned to prison for the commission of a violent crime within 14 months of parole. See JOHN MONAHAN & LAURENS WALKER, SOCIAL SCIENCE IN LAW: CASES AND MATERIALS 326-27 (3d ed. 1994). See generally MONAHAN, supra note 10. See also Eric S. Janus & Paul E. Meehl, Assessing the Legal Standard for Predictions of Dangerousness in Sex Offender Commitment Proceedings, 3 PSYCHOL. PUB. POL'Y & L. 33, 49 (1997)(reporting on actuarial prediction methods with 70-75% accuracy rates).

11. Compare ALLEN J. BECK, BUREAU OF JUSTICE STATISTICS, RECIDIVISM OF PRISONERS RELEASED IN 1983 5 tbl.9 (1989) [hereinafter BECK STUDY] (reproduced as Table 1, infra) with BUREAU OF JUSTICE STATISTICS SOURCEBOOK (1993) at 423 tbl.4.2 [hereinafter BJS SOURCEBOOK]. The latter Table indicates a national arrest rate of 5,648 per 100,000 (5.6%) in 1991 for all causes, including driving under the influence, drunkenness, and disorderly conduct. The rate for violent and nonviolent "index crimes" was 1,199 per 100,000 (1.2%). Table 9 of the Beck study reported a three-year recidivism rate for released prisoners of 62.5% for all charges, and a three-year rate of 22.7% for violent offenses. (Note that the arrest rate in Table 4.2 of the BJS SOURCEBOOK includes women and children. The arrest rate for adult males is higher. For example, in 1991, the homicide arrest rate for males aged 25-29 was 31.2 per 100,000. See BJS SOURCEBOOK, supra, at 449 tbl. 4.16. The homicide arrest rate in the general population in 1991 was 9.8 per 100,000. See id. at 423 tbl. 4.2. (I have chosen the 25-29 age group because the average age of the released prisoners in the Beck study was 27.) Even making this adjustment, the arrest rate for the released prisoners is many times that of males in the same age group.) Released prisoners also have high conviction rates. See BECK STUDY, supra, at tbl.8; STEPHEN B. KLEIN & MICHAEL N. CAGGIANO, BUREAU OF JUSTICE STATISTICS, THE PREVALENCE, PREDICTABILITY, AND POLICY IMPLICATIONS OF RECIDIVISM 10 tbl.2.2 (1986). Of course, the figures are affected to some extent by police selectivity, that is, by police focus on persons known to have records. Self-report studies also, however, indicate continuity of offending. See DAVID P. FARRINGTON, FURTHER ANALYSES OF A LONGITUDINAL STUDY OF CRIME AND DELINQUENCY. FINAL REPORT TO NATIONAL INSTITUTE OF JUSTICE (1983), cited in DAVID P. FARRINGTON ET. AL., UNDERSTANDING AND CONTROLLING CRIME 48 (1986).
is in issue, propensity evidence can be used in conjunction with incident-specific evidence in determining a question of historical fact about a particular occurrence. When evidence shows a person's high comparative propensity to engage in low base rate behavior, there is a great difference between predicting what might happen and finding out what did happen. When determining what did happen, evidence about a person's propensity can link up with incident-specific evidence in a way that makes it highly probable that the person committed an unusual act. That can be the case even if the same propensity evidence, by itself, would not justify a prediction that in the future the person would commit that unusual act. For example:

(1) Suppose that M is having an affair with F. One cannot predict that M will more likely than not kill F's spouse; very few lovers kill their rivals. But if we know that F's spouse has been murdered by someone using M's knife, then the evidence of the affair can become quite useful in deciding whether M was the murderer. While the propensity of lovers to kill their rivals is low in absolute terms, it is high in comparison with the propensity of persons who have no discernible motive to kill.

(2) While it may be true that only one in a thousand men who physically abuse their wives go on to kill them, if a wife is found murdered and the husband is a suspect, the evidence of abuse is certainly worth considering.

(3) Similarly, we cannot say that someone who committed one murder will more likely than not commit another. But if we suspect D of murdering V, and in looking for V's body in D's basement we find the body of X, we should attend to that information carefully, even if the risk that one homicide will follow another is generally quite small. The same could be said if, in looking into D's back-

---

12. In this paper, "propensity evidence" is a category that includes but is not limited to character evidence. It includes evidence of narrow propensities, e.g., the propensity to give a left turn signal or to beat a particular spouse, that are not considered character in the eyes of the law. Of my three numbered examples in the text following the call to this footnote, only the third would (probably) be considered "character evidence" for evidence law purposes.

13. In response to an argument by Alan Dershowitz during a television appearance that evidence of spousal battering was "massively irrelevant" in a murder case because only one in a thousand batterers goes on to kill, the statistician I.J. Good, assuming Dershowitz's numbers to be accurate and using data about the incidence of homicide from the World Almanac, calculated that the probability is one-third to one-half that a battering husband murdered his wife, given that we have no information other than that he battered her and that she was murdered. See I.J. Good, When Batterer Turns Murderer, 375 Nature 541 (1995). Upon learning that Dershowitz's statistic referred to the probability of murder per assault and not to the batterer's lifetime probability of murder, Good recalculated and raised his estimate to 0.9. See I.J. Good, When Batterer Becomes Murderer, 381 Nature 481 (1996).
ground, we discover that he committed a dissimilar murder in another state.14

Where conduct has a very low base rate in the general population, prediction that those who have done it before will do it again produces many false positives. No one engages in the conduct very often—not even actors with an unusually high propensity.15 But the reasons for caution begin to erode when it is undeniable that someone committed the low base rate act and incident-specific evidence makes a particular actor a suspect. Then, a history of similar extreme conduct (homicide, for example) can be highly incriminating—all the more so because of the unusual nature of the past acts.16

14. The three examples given all involve some form of propensity evidence, but only the third would be labeled “character” by jurists. The first example would be considered “motive” evidence. The second is closer to character than the first, but in spouse-murder cases evidence of prior spousal abuse is generally admitted under a non-character theory, one of which could be motive, e.g., that the accused was motivated by animosity toward a particular person, as exemplified by the prior abuse. For discussion and citation of authority, see Roger C. Park, Character Evidence Issues in the O.J. Simpson Case—Or, Rationales of the Character Evidence Ban, with Illustrations from the Simpson Case, 67 U. COLO. L. REV. 747, 752-54 (1996). Nonetheless, the principle being illustrated is the same in all three examples.

15. Probably the most common and surely the most significant error made by clinicians in predicting violent behavior is the ignoring of information regarding the statistical base rate of violence in the population in question . . . . For at least the past 25 years it has been known that it is virtually impossible to predict any ‘low base rate’ event without at the same time erroneously pointing the finger at many ‘false positives.’

MONAHAN, supra note 10, at 33.

Assume that one person out of a thousand will kill. Assume also that an exceptionally accurate test is created which differentiates with ninety-five percent effectiveness those who will kill from those who will not. If 100,000 people were tested, out of the 100 who would kill 95 would be isolated. Unfortunately, out of the 99,900 who would not kill, 4995 people would also be isolated as potential killers. In these circumstances, it is clear that we could not justify incarcerating all 5,090 people. If, in the criminal law, it is better that ten guilty men go free than that one innocent man suffer, how can we say in the civil commitment area that it is better that fifty-four harmless people be incarcerated lest one dangerous man be free?

Joseph Livermore et al., On the Justifications for Civil Commitment, 117 U. PA. L. REV. 75, 84 (1968). Compare the application of the hypothetical test to suspects who, by incident-specific evidence independent of that used for the test, were known to be 50% likely to have committed murder. In such a situation the test would raise the probability to 95%; i.e., it would raise the odds from 1 to 1, to 19 to 1. See also LEONARD BERKOWITZ, AGGRESSION: ITS CAUSES, CONSEQUENCES, AND CONTROL 286-87 (1993) (prediction of homicide will often be wrong because of low base rate). Cf. LEE ROSS & RICHARD E. NISBETT, THE PERSON AND THE SITUATION: PERSPECTIVES OF SOCIAL PSYCHOLOGY 134-35 (1991) (examining errors in lay reasoning about dispositions and noting that “[t]he costs of going against the base rate were particularly dramatic when the base rates were extreme”).

At the guilt phase of a criminal proceeding, the rules of evidence purport to exclude other-crimes evidence when offered for character inferences. At the punishment phase, other-crimes evidence is broadly admissible. Perhaps the law should be more confident in the value of other-crimes evidence as circumstantial evidence of guilt (when taken with other evidence), and less confident in its value in identifying persons who should be imprisoned for long terms. A system of short sentences with free use of probative evidence would show greater respect for human dignity than a system that sometimes excludes other-crimes evidence in determining guilt, but which then imposes Draconian sentences on prior offenders who claim the right to trial. Moreover, the fact-finder can use other-crimes evidence for its probative value in showing similar conduct by an actor soon after the other crimes. When used in sentencing for purposes of prediction, such evidence is being used partly to predict behavior of the defendant years after the crime that led to the sentence. The aging process has a powerful effect in reducing criminal propensity, making such predictions dicey.

(i) Two Illustrations of Probative Value

I will try to illustrate the relevance of prior crimes with two examples of situations in which a single instance of behavior shows an unusual propensity to commit a crime.

17. Only in exceptional circumstances do the Sentencing Guidelines limit consideration of prior criminal conduct. See, e.g., U.S. SENTENCING GUIDELINES MANUAL § 4A1.2(e) (1995) (sentence imposed more than ten years ago is not counted).

18. Of course sentencing has other purposes, such as retribution and general deterrence. But to the extent that it uses other-crimes evidence to lengthen sentences, it seems aimed mainly at preventive detention. The defendant has already paid for his prior crimes, and distinguishing between offenders who have records and those who do not has a somewhat indirect and speculative impact on general deterrence. For example, the Federal Sentencing Guidelines require that a judge “determine the defendant’s criminal history category” at the time of sentencing. U.S. SENTENCING GUIDELINES MANUAL § 1B1.1(f) (1997). The category is determined by a formula in which a judge assigns points for each prior sentence of imprisonment received by the defendant, more points if the charged offense was committed while on probation, parole, work release, etc., and still more points if the offense was committed within two years of release from a previous imprisonment. See id. § 4A1.1(a-f). The explanation for this formula contained in the introductory commentary to the Criminal History chapter of the rules is that “repeated criminal behavior is an indicator of a limited likelihood of successful rehabilitation.” Id. at § 4A.

19. For a powerful presentation of the evidence on age effect, see GOTTFREDSON & HIRSCHI, supra note 10, at 123-44.

20. The illustration is inspired by Richard O. Lempert, Modeling Relevance, 75 MICH. L. REV. 1021 (1977), though there is no guarantee that he would approve of my method or conclusion.
Drunk Driving Deaths

My first example seems virtually immune to selection bias by officials enforcing the criminal law. In their study of traffic deaths in North Carolina, Brewer and his colleagues\(^2\) found that 45% of the 7499 drivers killed in fatal accidents during the study period had blood alcohol levels above the specified levels. They then looked at the traffic records of those drivers in the same state for the preceding five years. Drivers with elevated blood alcohol at the time of death were 11 times as likely to have one or more North Carolina arrests for driving while impaired than were the other drivers.\(^2\) If all we could know about a driver was whether he died in a fatal accident while driving in North Carolina in that period and whether he had a prior arrest for DUI in North Carolina, then: (1) upon learning that a driver died while driving we should believe there to be a 45% probability that he had an elevated blood alcohol level at the time of death, or odds of a bit less than 1 to 1; and (2) upon also learning that the driver had a prior DUI arrest we should raise that probability estimate to 90%, or odds of 9 to 1.\(^2\)

Of course that is not all we can know. In an actual case, we would be suspicious if the party seeking to prove elevated blood alcohol did not produce the results of a blood or breath test. But if more desirable alternatives are not available, the prior record evidence would tell us something well worth our attention.

Homicide Recidivism

Homicide is a very rare crime, hence we should be reluctant to

\[^{21}\text{See Robert D. Brewer et al., The Risk of Dying in Alcohol-Related Automobile Crashes Among Habitual Drunk Drivers, 331 New Eng. J. Med. 513 (1994).}
\[^{22}\text{See id. at 514.}
\[^{23}\text{Using the odds form of Bayes' theorem, if P(G) is the prior probability of elevated blood alcohol level; P(EIG) is the probability of finding a prior DUI arrest, given elevated BAC; and P(GIE) is the probability of elevated BAC, given the evidence of a prior DUI arrest:}
\]

<table>
<thead>
<tr>
<th>Prior odds:</th>
<th>Likelihood ratio:</th>
<th>Posterior odds:</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(G)</td>
<td>P(EIG)</td>
<td>P(GIE)</td>
</tr>
<tr>
<td>P(-G)</td>
<td>P(EI-G)</td>
<td>P(-GIE)</td>
</tr>
</tbody>
</table>

\[X = \frac{P(G) \cdot P(EIG)}{P(-G) \cdot P(EI-G)} = \frac{45}{55} \cdot \frac{11.1}{1} = 90.08\%\]
predict that a murderer will do it again. But comparative propensity can be high even when base rates are low. A comprehensive Bureau of Justice Statistics study indicates that a prisoner released from a homicide sentence is almost five times as likely to commit homicide as a prisoner released after serving a term for another serious crime, despite the fact that (a) persons imprisoned for homicide are likely to have served long sentences and hence be older than other prisoners when released, and (b) the sample did not include homicide offenders who were executed or held for life. This number suggests an impressive comparative propensity, which when combined with independent evidence of homicide, could be highly probative. As might be expected, the comparative propensity is dramatically higher when one compares homicide offenders with the general population rather than with other offenders.

24. Same-offense re-offending is lower for homicide than for other offenses (a Bureau of Justice Statistics study indicates that 6.6% of released murderers will be rearrested for homicide within three years, a smaller number than for the other crimes studied). See BECK STUDY, supra note 11, at 5 tbl.9 (reproduced as Table 1, infra).

25. Id.

26. The average age of the prisoners released in the Beck study was 27. Presumably the released murderers were older, having served longer sentences. Even 27 is past the prime age for the crime. See GOTTFREDSON & HIRSCHI, supra note 10; BJS SOURCEBOOK, supra note 11, at 449 tbl.4.17 (showing dramatic age effect for homicide arrests).

27. If one derives an annual recidivism rate by dividing the three-year rate reported by the Beck study by three, then the arrest rate for homicide among prisoners released from a sentence of homicide is 2.2%, or 2200 per 100,000. The arrest rate in the general population for homicide was 9.8 per 100,000 in 1991. See BJS SOURCEBOOK, supra note 11, at 423 tbl.4.2. Among males in the prime age group of 18-20, it was 41.8 per 100,000 in 1985 (the end of the period covered by the Beck study). See id. at 449 tbl.4.16. The comparison is imperfect, however, because the Beck study uses a more inclusive definition of homicide (murder, nonnegligent homicide, negligent homicide) than that used in the BJS statistics for the general population (murder, nonnegligent homicide). Comparing the arrest rate for any violent crime of homicide offenders (21.5% in three years, 7.2% per year) with that of the general population, one gets a ratio of 25 to 1. Though it is somewhat speculative to correct the homicide figures, my guess is that the ratio is somewhat between the 25 to 1 for all violent crimes and the 224 to 1 for homicides, and closer to the latter. The ratio would be higher if one considered only the first year after release, since released prisoners commit more crimes during that period than in subsequent periods. See BECK STUDY, supra note 11, at 7.

The comparative propensity estimates are probably inflated to some extent by the higher scrutiny given to prior offenders by police. However, in homicide cases (a) there is usually no doubt that the crime occurred—homicide incidence figures are regarded as among the most trustworthy of those available in the study of crime, and (b) most homicide arrests are referred for prosecution, suggesting that they are not random roundups. See, e.g., DIV. OF CRIMINAL JUSTICE INFORMATION SERVICES, CALIFORNIA DEPT OF JUSTICE, HOMICIDE IN CALIFORNIA 90 tbl.35 (1995) (93% referred in 1995; 80% in 1986). Of course it is possible that police are solving a large percentage of homicides by falsely pinning them on past offenders, but the dangers of such an approach seem likely to
In the 1970s and 80s, several defenders of the character evidence ban found support in a theory of personality that is sometimes labeled as situationism or the theory of specificity. These legal scholars argued that human conduct is determined by the pressures of particular situations, not by enduring cross-situational character traits, and that one cannot predict (or postdict) behavior by reference to concepts of human propensity. This argument about the unpredictability of human behavior was often stated quite strongly, even to the point of saying that character evidence was irrelevant.

Recently a more cautious attitude has become apparent, which is perhaps more favorable toward reception of evidence of past mis-

limit its prevalence. Arrest data as a measure of the underlying offense compares well in terms of validity with other social science data. See Joseph G. Weis, Issues in the Measurement of Criminal Careers, in 2 CRIMINAL CAREERS AND "CAREER CRIMINALS" 1, 13-14 & n.11 (Alfred Blumstein et al. eds., 1986). Finally, the dramatic decrease in homicide arrests in the post-teen years, see BJS SOURCEBOOK. supra note 11, at 449 tbl. 4.17, suggests that there must be some limit on the degree to which police routinely “round up the usual suspects.” Otherwise, one would expect persons in late youth or early middle age to compromise a larger proportion of homicide arrestees, as a consequence of police round-ups of old offenders with records of violence.

28. For example, Professor Spector wrote:

This theory [of specificity] says that character traits such as honesty or carefulness do not exist. How a person behaves depends on the specific situation in which that person finds himself . . . . The implications for evidence law are clear. Character evidence used as a basis for predicting human behavior is useless . . . . The legal conclusion therefore ought to be that evidence of character has no probative value.


Noting that Gordon Allport, a leading trait theorist, had written that his earlier views seemed to neglect the variability induced by ecological, social, and situational factors, Professor Lawson concluded that “the theory of behavior that was so compatible with the law’s notions about character has ceased to have any scientific recognition.” Robert G. Lawson, Credibility and Character: A Different Look at an Interminable Problem, 50 NOTRE DAME LAWYER 758, 783 (1975). Professor Lawson did, however, note a body of scholarship that viewed personal characteristics as interacting with situations to produce conduct, and took a less absolute view of the power of situations than that taken by Professor Spector; however, he concluded that “the data most certainly reveal that the probative worth of character evidence is very low.” Id. at 786.

conduct. Susan M. Davies, examining the literature on trait theory, situationism, and interactionism, has drawn from it the lesson that evidence of past misconduct can indeed be useful, at least if one takes into account factors such as similarity between surrounding circumstances, the number of prior instances, proximity in time, and whether the conduct is "inherently likely to be repeated." 29 Professor Miguel Méndez, one of the pioneers in the examination of personality theory for insights about character evidence, has concluded from a review of recent research that its implications may now be more favorable to the reception of predisposition evidence than in the 1980s. 30 I have an even more cautious view of the significance of the studies of specificity of behavior for law reform, simply because the issues of behavioral consistency and stability addressed by these studies are different from the issues likely to be addressed by reformers of the law of character evidence.

I will start by distinguishing three issues in the law of character evidence. The first is whether witnesses should be allowed or required to testify in terms of generalizations about character—for example, by giving reputation or opinion testimony about a defendant's general traits of honesty, lawfulness, or peaceability. 31 The second is whether the law should allow impeachment of the credibility of defendants with evidence of prior convictions, while simultaneously instructing the jury to use the evidence, not as evidence of propensity to commit the substantive crime, but as evidence bearing only on credibility. 32 The third is whether specific-act evidence that the defendant has committed other similar crimes should be admitted on the issue whether the defendant committed the crime charged. 33


If, as Mischel found in the sixties, even trivial situational differences may reduce the expected correspondence or correlation to zero, then the law was justified in banning the use of predisposition evidence. Under this view, consciously cheating clients on sales probably would tell us nothing about whether the defendant consciously understated his income or overstated his deductions. Mischel and Shoda's new work challenges the prevailing scientific and legal conclusions about the value of predisposition evidence. If they are right about the relative stability and invariance of the basic personality structure and about the stable patterns of variability flowing from that structure, then consciously cheating clients might indeed tell us something about whether the defendant consciously evaded his tax obligations.

Id. at 234.
31. See FED. R. EVID. 405(a).
32. See FED. R. EVID. 609.
33. See FED. R. EVID. 404(b), 413-14.
To be fair to legal scholars who found lessons for the law of character evidence in personality theory, I should note that they were primarily concerned with the relevance of the literature to the first two issues, especially to the use of prior convictions to impeach a criminal defendant. There, where the situational pressures are strong and the relevance of general disposition weak, even a weak version of situationism would help us reach the conclusion that the law’s treatment of the problem is wrong.\(^{34}\) In addition, social psychology has much to say about whether jurors will follow the odd limiting instructions that attempt to prevent substantive use of impeachment evidence.\(^{35}\)

Psychology scholarship has less to say about the third issue, which is more salient in recent controversies about character evidence because of issues raised by legislation allowing other instances of sexual assault to be admitted in a sexual assault case and other instances of child molestation to be admitted in a child molestation case.\(^{36}\) Taking that legislation as a model, a law reformer with abolitionist sympathies might suggest that this approach be extended to other crimes—that other homicides be admissible in homicide cases, or more broadly that other acts of criminal violence be admissible when criminal violence is charged.

On this issue the literature from social psychology and personality psychology does not point strongly to one conclusion or the other. Consider the two books most heavily relied upon in law review articles, the Hartshorne and May study\(^ {37}\) and Walter Mischel’s *Personal-ity and Assessment*.\(^ {38}\) These are seminal works, respected and influential; but they were not, of course, produced for the purpose of answering questions about evidence law. In assessing their applicability to issues of evidence law, legal scholars should consider the behavior examined in those works.

\(^{34}\) This issue was the primary focus of Professors Spector and Lawson, cited at note 28, *supra*. For an example of a scholar who made a strong case against use of convictions to impeach without overt reference to psychology, see H. Richard Uviller, *Evidence of Character to Prove Conduct: Illusion, Illogic, and Injustice in the Courtroom*, 130 U. P.A. L. REV. 845 (1982); for a scholar who reached the same conclusion without the necessity of a strong version of situationism, see Richard Friedman, *Character Impeachment Evidence: Psycho-Bayesian[!] Analysis and a Proposed Overhaul*, 38 UCLA L. REV. 637 (1991).


\(^{36}\) See *FED. R. EVID.* 413-15.


\(^{38}\) WALTER MISCHEL, *PERSONALITY AND ASSESSMENT* (1968).
The Hartshorne and May work presented a series of studies of honest and deceptive behavior of schoolchildren. For example, under observed conditions, children were given the opportunity to steal coins that they had used for games and counting exercises, to lie in order to avoid disapproval or win approval, or to cheat while self-scoring a test. Hartshorne and May found that dishonest behavior in one situation was only modestly correlated to dishonest behavior in dissimilar situations. Even as applied to conclusions about dishonest behavior, one might object that the study involved children, who may not have developed as much consistency in moral behavior as adults. More importantly, to the evidence scholar interested in whether crimes of violence should be admitted as propensity evidence, it is a long jump from a study showing that children are inconsistent in behavior that involves no obvious harm to others or danger to themselves, to the conclusion that acts of criminal violence have no predictive value.

Mischel examined an extensive body of literature and many empirical studies, but the literature did not include studies of criminal violence. He drew on the Hartshorne and May work as a study of "moral" behavior. He also examined an extensive body of literature

39. See HARTSHORNE & MAY, supra note 37, at 90-94.
40. See id. at 94-97.
41. See id. at 98-103.
42. See id. at 51-55.
43. While the average correlations among deception tests of the same type ranged from .440 to .836, the average correlations between single tests of different techniques ranged from only -.003 to .312. See id. at 382-83. The authors stated that "as we progressively change the situation we progressively lower the correlations between the test.... [W]e interpret these facts to mean that the consistency of the individual is a function of the situation." Id. at 384. Hartshorne and May concluded that "[t]he results of these studies show that neither deceit nor its opposite, 'honesty,' are unified character traits, but rather specific functions of life situations." Id. at 411.
44. See Friedman, supra note 34, at 650 n.40.
45. Mischel did refer briefly to criminal behavior in the course of making a comparison that, if anything, seems to look favorably on the predictive value of evidence of a past criminal history (at least in comparison with predictions based on psychological testing and clinical assessment). The reference occurs in Chapter 5, in which Mischel examines the utility of personality tests and clinical judgments used for global trait attributions (and sometimes for prediction of behavior). See MISCHEL, supra note 38, at 103-48. As an example, he quotes a clinical report predicting that a patient's passive-aggressive character will cause him to be rejected by long-term sexual partners. See id. at 104-05. In the course of examining the utility of trait attributions based on clinical judgment and testing, Mischel indicates that actuarial measurements based on past behavior fare relatively well as predictors, noting that in one study of predicting parole violations, a "base expectancy index" created from a "simple record of past behavior" (prior commitment, frequency of previous delinquencies, type of offense, age at first admission) was found to be the "best single predictor... excelling all personality scales." Id. at 136.
46. See id. at 23-25.
studying children’s ability to delay gratification under varying lab conditions;47 the generality of such traits such as dependency,48 extroversion,49 rigidity,50 and tolerance for ambiguity,51 cognitive avoidance of anxiety-provoking cues;52 conformity;53 and conditionability.54 Again, the underlying studies did not deal with criminal or violent behavior. Even for the behavior studied, Mischel did not offer extreme claims about the absence of individual differences or the irrelevance of prior behavior. He even looked favorably upon the use of other act evidence for prediction in certain situations, considering it superior to a clinician’s psychodynamic assessments using concepts such as “infantile dependency needs,” and “passive-aggressive character make-up.”55

47. See id. at 81-83. For example, researchers studied how situational constraints influenced children’s desire to wait for a bigger reward instead of taking a smaller one right away.
48. See id. at 27-28. An example of a dependency study examined by Mischel is Robert Sears, Dependency Motivation, in NEBRASKA SYMPOSIUM ON MOTIVATION 25, 25-64 (M.R. Jones ed., 1963). Sears studied the cross-situational consistency of different dependency behaviors in nursery school children—such as touching a teacher or child, requesting reassurance, and seeking attention. See MISCHEL, supra note 38, at 27.
49. See MISCHEL, supra note 38, at 100.
50. See id. at 29.
51. See id. at 28.
52. See id. at 30-32.
53. See id. at 29.
54. See id. at 32.
55. For an example of the type of psychodynamic assessment about which Mischel was dubious, see id. at 104-05.
Mischel himself has complained about oversimplified and exaggerated interpretations of his ideas.

[M]any of the original issues in the challenge to global dispositional approaches to personality also have tended to become oversimplified in efforts to recollect and summarize them. Thus, for example, the critique in PERSONALITY AND ASSESSMENT (Mischel, 1968) has not infrequently been viewed as a broadside attack on personality itself and even as an effort to replace dispositions and indeed people with situations and environments as units of study.


Mischel has also stated that he was reacting against the commonplace practice among clinical psychologists
to attempt extensive decision making about a person’s life and future on the basis of a relatively limited sampling of personological ‘signs’ or ‘trait indicators.’ As a clinically trained psychologist . . . I was especially sensitized to the potential hazards of such attributions, of such categorizations too often made on the basis of scanty evidence. Stimulated by my mentors, my hope was to call attention to the specific reciprocal interactions between person and context and hence to the need to examine those interactions in fine-grain detail when attempting person-centered decisions and predictions. My concern was that clinicians, like other scientists and indeed like the ordinary layperson, try to infer, generalize, and
A 1991 text by Ross and Nisbett surveys these studies and other more recent ones on the specificity of behavior. An examination of their work confirms the impression that, in the disciplines of personality and social psychology, conclusions about cross-situational consistency of behavior have been based on studies that examine non-violent, noncriminal behavior. By this observation I mean no predict too much from too few observations. As has become increasingly recognized, the judgments of clinicians—like everyone else’s judgments—are subject to certain systematic biases in information processing that can produce serious distortion and oversimplifications in inferences and predictions. I was sensitive to the fact that clients—like other people—don’t describe themselves with operational definitions. They invoke motives, traits, and other dispositions as ways of describing and explaining their experiences and themselves. Much of the assessor’s task, it seemed to me, should be to help people in the search for such referents for their own personal constructs, instead of forcing the assessor’s favorite dispositional labels on them.

My reading of the available data in 1968 did not imply that useful predictions cannot be generated, nor that different people will not act differently with some consistency in different classes of situations. But in my view the data did then and do now imply that the particular classes of conditions must (1) be taken into account far more carefully than in the past, (2) tend to be much narrower and subtler than traditional trait theories have assumed (especially when one goes beyond recall-based ratings), and (3) require, for purposes of important individual decision-making or for specific predictions, highly individualized assessments of subjective meanings because behavior tends to be idiosyncratically organized within each person on his or her personal dimensions of similarity that do not necessarily match the categories of the trait psychologists. To me the data then and now also suggest that the inferences about global underlying traits and dispositions—especially, indirect inferences—often have less utility for most assessment efforts to predict or therapeutically modify individual behavior than do more economical, alternative analyses based on more direct data such as people’s past behavior in similar situations or their direct self-reports and self-statements.

Id. at 280-81 (citations omitted). See also Mischel, supra note 38, at 281-82 (“[M]ost trait and state theories assume internalized behavioral dispositions relatively independent of stimulus conditions. It is the existence of such stimulus-free, highly generalized behavioral sets—not the occurrence of long-term individual differences in response to stimuli—that is unsupported by the data reviewed in earlier chapters”); 282-301 (acknowledging personality consistencies in natural environments). In recent work, Mischel has written that there are “enduring individual differences in the features of situations that individuals select and the cognitive-affective mediating units (such as encodings and affects) that become activated, and that interact with and activate other mediating units (e.g., expectancies, goals, behavioral scripts and plans) in the personality system.” Walter Mischel and Yuichi Shoda, A Cognitive-Affective System Theory of Personality: Reconceptualizing Situations, Dispositions, Dynamics, and Invariance in Personality Structure, 102 PSYCHOLOGICAL REVIEW 246, 246 (1995). The passage quoted is fully consistent with the view that, for example, some men are far more likely than others to interpret ambiguous conduct as inviting sexual aggression or to believe and act on rape myths.

57. In Chapter 2, “The Power of the Situation,” Ross and Nisbett describe “some of
the classic studies of social psychology" to illustrate "the power of situational factors." Id. at 25. They begin by discussing several laboratory studies, starting with the early work of Muzafer Sherif, and later, Solomon Asch, that demonstrate the conformity pressures exerted by groups. See id. at 28-32. The studies required subjects to make judgments based on various types of stimuli, including the apparent movement of a light dot, the relative length of several comparison lines to a standard line, the relative pitch and length of comparison tones to a standard tone, mathematical problems, general-knowledge items, and social and political judgments. See id. at 28-35. Ross and Nisbett next describe the Bennington Study conducted in the 1930s by Theodore Newcomb. See id. at 35. Newcomb considered the political attitudes of the generally conservative young women who entered Bennington College, and found that the women's views and preferences shifted far to the left within a few years of exposure to the "Bennington milieu." See id. Ross and Nisbett then present studies of intergroup relations conducted by Sherif and several colleagues, who undertook field experiments at a summer camp where 12-year-old boys were separated into two cabins and later placed in competitive and cooperative situations. and by Henry Tajfel and his colleagues, who sought to demonstrate that merely categorizing subjects into different groups could produce favoritism and discrimination between the different group members. See id. at 38-41. They also describe bystander intervention studies that placed subjects in situations where they could come to the aid of a person in distress or report a staged emergency. See id. at 41-44 (describing studies beginning with the seminal work of John Darley and Bibb Latané). Ross and Nisbett then present studies of intergroup relations conducted by Sherif and several colleagues, who undertook field experiments at a summer camp where 12-year-old boys were separated into two cabins and later placed in competitive and cooperative situations, and by Henry Tajfel and his colleagues, who sought to demonstrate that merely categorizing subjects into different groups could produce favoritism and discrimination between the different group members. See id. at 38-41. They also describe bystander intervention studies that placed subjects in situations where they could come to the aid of a person in distress or report a staged emergency. See id. at 41-44 (describing studies beginning with the seminal work of John Darley and Bibb Latané). Ross and Nisbett conclude their discussion of the power of circumstances to produce behavior by reviewing various studies that attempt to explain the power of societal and group influences over dissenting individuals. See id. at 44-46. Next, they discuss studies demonstrating that "small differences between situations often are associated with very large behavioral differences." Id. at 46. For example, one study showed that the United States government was able to dramatically increase sales of war bonds by changing its advertising from general to specific appeals to the public. See id. at 47-48. Another study considered the effect of being in a hurry on the subject's willingness to be a "Good Samaritan" and aid a person in distress. See id. at 48-50 (reviewing studies of various situational determinants on altruistic behavior). Yet another line of studies considered the effect of eliciting minimal compliance, or getting a "foot in the door," on subsequent requests—for example, the effect on a subject's willingness to donate money to the Cancer Society after previously agreeing to wear a lapel pin advertising the fund drive. See id. at 50-52. Ross and Nisbett conclude their chapter on the power of the situation by providing a detailed analysis of Milgram's studies on obedience as demonstrated by the willingness of subjects to administer ever-increasing electric shocks to another person in a supposed educational experiment. See id. at 52-58. The latter experiments (Milgram) might seem to measure conduct similar to criminal violence, but they are more appropriately viewed as studies of obedience to authority rather than exhibitions of free-lance violence. See Michael R. Gottfredson and Travis Hirschi, A Control Theory Interpretation of Psychological Research on Aggression, in AGGRESSION AND VIOLENCE: SOCIAL INTERACTIONIST PERSPECTIVES 47, 53 (Richard B. Felson and James T. Tedeschi eds., 1993). Gottfredson and Hirschi agree with Milgram that the "subject's shocking behavior has nothing to do with aggression," and they suggest that such studies "in fact measure compliance, obedience, or... acquiescence." Id. Furthermore, Gottfredson and Hirschi demonstrate that laboratory measures of aggression have a negative correlation with official records of delinquency and with self-report of theft, violence, and drug use, concluding that "the more 'aggressive' the subjects, the less likely they are to commit delinquent acts." Id. at 55.

In Chapter 4, "The Search for Personal Consistency," Ross and Nisbett "review evidence concerning the explanatory and predictive power of personality traits such as extroversion, honesty, and dependency," or rather "their apparent lack of power." ROSS &
criticism of the field, but simply note that law professors should be cautious in generalizing that its findings apply to other-crimes evidence. Perhaps because the consistency of violent criminal behavior is not easy to study using the tools of social psychology, these works do not focus on that type of behavior.

Is there a reason to think that criminal violence might be a more consistent and stable type of behavior than the behavior studied in these works? Some clues to the answer may be found in the studies by social and personality psychologists of “aggression,” a somewhat similar (but much broader) category of behavior. These studies report a relatively high degree of consistency and stability. Another, NISBETT, supra note 15, at 91. They note Mischel’s role in the “challenge of 1968” to the conventionally held theory that human behavior was predictable across different situations. See id. at 95-96. Their review of the types of studies Mischel cited indicates that they did not concern criminal or violent conduct. The review describes the work of Theodore Newcomb, who studied “problem” adolescent boys at a summer camp to examine the consistency of extroverted behavior—whether a boy who was talkative at lunch would also disrupt the “quiet hour.” See id. at 96-97. The review also mentions the Hartshorne and May study of deceptive behavior in schoolchildren, and the study by Robert Sears on the dependency of nursery school children. See id. at 98-100. Finally, Ross and Nisbett note Mischel’s studies on children’s willingness to delay gratification as a demonstration of the weakness of cross-situational consistency of behavior and the power of situational factors. See id. at 101-02.

58. One important tool of social psychology, the controlled laboratory experiment, is simply unusable.

59. Gottfredson and Hirschi question whether the results of lab experiments in which the subject believes he is shocking a confederate of the experimenter correlate with the type of aggression involved in delinquent behavior. See Gottfredson & Hirschi, supra note 57, at 53-55. The criticism does not, however, apply to other measures of aggression, such as peer ratings and official reports.

60. For example, in his Introduction to Personality text, Mischel writes: Aggression is a dimension of behavior on which stable individual differences have been identified beginning in grade school. These differences remain stable even over long periods of time. In longitudinal studies of adolescent boys in Sweden, for example, peers rated each others’ aggressiveness on a number of different rating scales, such as verbal aggressiveness in response to criticism from teachers, and how easily the child starts fights at school. Most boys who were seen as aggressive in the sixth grade were still seen as aggressive in the ninth grade, even when the specific classes and peers changed. Similar results come from studies that span long time periods, linking measure of aggression at the age of eight years to follow-up measures of aggression more than two decades later. For example, individuals who are rated high on aggression at age eight tend to be aggressive and more likely to commit antisocial acts when they are thirty years old.

MISCHEL, INTRODUCTION TO PERSONALITY 480-81 (5th ed. 1993) (citations omitted). See also BERKOWITZ, supra note 15, at 131 (extreme aggressiveness—an “unusually great readiness to hit others physically and verbally”—linked to “other antisocial tendencies”); GOTTFREDSON & HIRSCHI, supra note 10, at 50 (aggression scores comparable in stability to intelligence scores; “what is not arguable is that aggressive behavior, however engendered, once established, remains remarkably stable across time, situation, and even
stronger clue is found in the work of criminologists and other scholars of crime, whose view of individual differences is quite different from that of the legal scholars who have used the literature from social psychology. 61

Susan Davies’ influential article62 used data from social and personality psychology to argue that trait information could be useful in predicting behavior and that lay people could predict with significant accuracy. 63 I would go just a step further, and suggest that in our use of the studies from social and personality psychology, including ones cited by Davies, we need to heed Daubert’s warning about “fit.”64 We should not be too eager to generalize from studies of noncriminal behavior to conclusions about the probative value of other-crimes evidence. 65

Davies was responding to earlier legal scholars who applied the studies of the social and personality psychologists to the problem of character evidence. She argued persuasively that they overemphasized the power of situations and underestimated the predictive value of evidence of other acts, but she accepted their view of the applicability and cogency of the underlying body of personality research. If one adopts her perspective, one is likely to emphasize the importance of similarities in situations, goals, and motives, and the importance of observing repeated instances of prior conduct.

Proximity and similarity are important. However, single act evidence can sometimes be useful, even in the absence of a strong showing of similarity. The scholarship from social psychology and personality psychology only demonstrates that it is difficult to predict from one instance of ordinary behavior that another specific instance of ordinary behavior will occur. For example, the fact that a child was observed to cheat during lab task #1, designed to give her an opportunity to cheat without apparent danger of detection, helps little in predicting whether she will behave dishonestly when observed the
next day doing lab task #2, a dissimilar test of dishonesty. However, the proposition just stated could be true, and both of the following propositions could also be true:

(1) a single instance of extreme behavior (such as murder) can demonstrate that the actor has a high comparative propensity to commit another instance of extreme behavior of the same general type. That is, it is reasonable to believe that a person who has

66. See Hartshorne & May, supra note 37, at 382-83; Ross & Nisbett, supra note 15, at 129 ("It will be recalled from our review of the consistency evidence in the previous chapter that the same data sets that show low consistency across different types of behavior and situations sometimes show rather high stability within similar types of behavior and situations.").

67. See Ross & Nisbett, supra note 15, at 116. In a section entitled "The Relative Likelihood of Extreme Behaviors," Ross and Nisbett illustrate the point that even low correlations can be useful for prediction in some circumstances. They hypothesize a cross-situational consistency correlation of .16 in an extroversion study. Then they note that "a correlation of that level can prove to be quite useful for one type of prediction task . . . [T]his task involves identifying people who are relatively more or less likely than their peers to score at either extreme of the distribution." Id. Suppose, they suggest, that on a single randomly sampled occasion a subject (Harry) scored in the 98th percentile on extroversion based upon the commission of an extraordinarily extroverted act (he was observed at an office party wearing a lampshade and reciting ribald limericks). That observation, they report, makes Harry over twice as likely as the ordinary individual to be engaged in highly extroverted behavior (in the 98th percentile) when next observed (and over five times as likely as a person who scored in the bottom 2 percent on the single prior observation). Although the chance that Harry will be engaged in such behavior when randomly observed is still quite small, he has a high comparative propensity. See id. To the present author, this illustration by Ross and Nisbett suggests that when there is substantial but conflicting incident-specific evidence about whether Harry engaged in the behavior on an occasion of interest in litigation, then other-instance evidence could be quite helpful.

Ross and Nisbett also state:
Even on the basis of a correlation as low as .16 for individual response, we will be able to make some pointed predictions about the relative likelihood that particular types of extreme responses will be shown by various specific individuals. Indeed, knowing that a particular individual has displayed an extremely "high" response on even a single occasion makes it safe to conclude that the person's response on some other occasion is much more likely to be extremely high than extremely low.

Id. at 111. Ross and Nisbett suggest that predictions of this nature are useful for a variety of "screening" purposes where the goal is to minimize or maximize extreme behavior. See id. at 116. They do not consider its value in shedding light on questions of historical fact in conjunction with other evidence.

Mischel himself thought that predictions could sometimes be based upon relatively fragmentary history or even single acts:
Perhaps most impressive, we are finding highly significant continuities, linking the preschooler's delay time while waiting for a couple of marshmallows to his or her cognitive and social competence and school performance a dozen years later (Mischel, 1983). Sometimes even single acts can predict meaningful life outcomes! For instance, the number of seconds preschoolers delayed the first time they had a chance to do so in our 1968-1972 studies, regardless of the specific
committed one such act has a propensity to commit that type of act that is many times greater than the propensity of a person chosen at random; and

(2) that high comparative propensity to engage in extreme behavior could have great value, when used in conjunction with incident-specific evidence, in showing whether a person committed a single act of that type of behavior on a particular occasion.

b. Flaws in Human Reasoning That Might Cause Misuse of Character Evidence

So far I have referred only to the literature dealing with cross-situational consistency and stability. Another body of literature, bearing on errors in human reasoning, has also been used by legal scholars to bolster the argument that character evidence is prejudicial because lay fact-finders will be unable to reach accurate results with it. These scholars argue, for example, that fundamental attribution error causes human decision-makers to attribute too much importance to dispositions, and to overlook situational influences.68

This body of scholarship has since developed in a way that may undermine their arguments: recent studies seem to be more favorable to lay reasoning than those studies relied upon by legal scholars in the 1970s and 80s.69 Moreover, in assessing the legal significance of delay situation they encountered, significantly predicted their rated social competence as high school juniors and seniors in 1982 (r=.34, N=99, p<.001).

Mischel, supra note 55, at 277.

Far from suggesting universal absence of cross-situational consistency, or the need to sample a large variety of behavior, the above passage suggests that a single act can be predictive of conduct years after. Note also that inability to delay gratification (lack of self-control) is, if Gottfredson and Hirschi are correct, see Gottfredson & Hirschi, supra note 10, at 89, a trait highly linked to criminal conduct.

Another example by Mischel appears earlier in the same work:

Environments, properly analyzed, can enhance prediction. So can data about the person's directly relevant past behavior. "Unless the environmental maintaining conditions change markedly, past behavior tends to be the best predictor of future behavior in similar situations, even when the samples of past behavior are very fragmentary, as in gross indices of previous martial history, occupation, or hospitalization." (Mischel, 1968, p. 292). A vivid example of the value of the past for predicting the future is the fact that an individual's past record of maladjustment and psychiatric hospitalization is a good index of future maladjustment and hospitalization (Lasky et al., 1959). Thus, a correlation of .61 was found between the weight of the patient's file folder and the incidence of rehospitalization.

Mischel, supra note 55, at 281.

68. See Spector, supra note 28, at 351-52; Teree E. Foster, Rule 609(a) in the Civil Context: A Recommendation for Reform, 57 FORDHAM L. REV. 1, 33 (1988); Lawson, supra note 28, at 778.

69. See Davies, supra note 29, at 524-29. Cf. David C. Funder, Errors and Mistakes:
research about constructs such as fundamental attribution error, it is crucial once again to formulate the legal issue with precision. In the case of a proposed character evidence reform that would let in evidence of prior violent crimes to show commission of the crime charged, the legal question might be framed as follows: will the trier of fact be more accurate in determining whether the defendant committed the crime charged if the trier is given evidence about the defendant’s commission of other crimes at other times?

With regard to that legal question, one can make the following observations:

(1) The legal implications of the attribution error research are two-sided. Giving judges discretion to exclude evidence on grounds that it will prejudice the jury invites judges to act on global dispositional attributions backed up by very little information—for example, by deciding that the present jury will give too much weight to spousal abuse as evidence of murder, or that a central city jury can't be trusted with evidence of police racism and misconduct. There is certainly no guarantee that a judge will be immune from attribution error in making decisions, including a decision about what might prejudice the jury whose voir dire the judge has supervised.

(2) Research that shows that subjects tend to make dispositional attributions even when the subjects lack relevant data is two-sided in its legal implications. Does it mean that human decision-makers cannot be trusted with evidence of other conduct of the actor, or that they would do better if given more such evidence? Consider the well-known experiment in which subjects were asked to form a judgment about whether a debater favored Fidel Castro.70 Even if the subjects were told that the debater had no choice, subjects were more likely to attribute a pro-Castro attitude to the debater if the debater spoke in favor of Castro than if he spoke against Castro. First, as Funder has pointed out, the experimental results may have been misleading because the experimenter may, by merely asking for an attribution, have implied that information allowing a rational judg-

---

ment was being offered. But whether or not that is the case, it is not clear what message such experiments carry for the law of evidence. Suppose that the subjects had been charged in a trial with the task of determining whether the speaker did indeed agree with Castro. The fact that they might make dispositional attributions without any evidence does not logically entail that evidence of disposition should be excluded. Perhaps the accuracy of fact-finding would be improved by "admitting" additional evidence about the political disposition of the speaker, such as evidence of the speaker's political affiliations or ideology.

Fundamental attribution error research indicates that human reasoners tend not only to make dispositional judgments freely, but also to overestimate the power of dispositions. But the legal policy implications are unclear because the research seems to show that humans will make dispositional attributions and overestimate the power of dispositions whether or not they are given relevant data about the history of the person whose disposition is being judged. So while the research might support the idea that triers should be cautioned against making that mistake, it does not carry a clear lesson that criminal history evidence should be excluded.

In fact, jurists could find the opposite lesson in some of the literature. Ross and Nisbett describe the tendency of lay dispositionists to fall into "interview illusion"—the "assumption that one can learn a great deal of useful information about people's personalities from a brief get-acquainted interview." They point out the greater predictive power of personal history information—for example, an applicant's grades in school.

71. See Funder, supra note 69, at 80.
73. Id.
74. Ross and Nisbett describe the "illusion" as follows:

   The literature reviewed to this point [on lay overconfidence in predictions based on dispositions] is helpful in understanding what we have called the 'interview illusion' (Nisbett & Ross, 1980). This is the assumption that one can learn a great deal of useful information about people's personalities from a brief get-acquainted interview. This belief may be called an illusion because the best available evidence on the predictive validity of unstructured interviews for estimating future college or graduate school performance, or job performance by blue-collar or white-collar workers, or professional success by executives, lawyers, doctors, or research scientists, indicates that the relevant correlations [between the perceived effect of the interview and the actual performance] rarely exceed the .10 or .15 range. The majority of studies, in fact, produce correlations of .10 or less (Hunter & Hunter, 1984).

Id. at 136.
75. Ross and Nisbett state that high school grades are a reasonably good predictor of college grades, with validities in the range of .30 and .45—and that "subjects tend to un-
The American criminal trial, with criminal history evidence excluded, could be viewed as the legal analogue to “interview illusion.” Consider a consent defense rape case in which the complainant testifies that she was raped, the defendant claims she consented, there are no eyewitnesses other than the two principals, and the circumstantial evidence is equivocal. The idea that the trial should be conducted without any reference to other acts of the defendant is analogous to the idea that hiring should be done solely on the basis of interviews. Excluding other act evidence does not prevent jurors from making character attributions. It merely means that the character attributions will be based on trivial and misleading data. Attributions might be based on factors such as the attractiveness of the defendant and complainant, the coherence of their stories (which have been polished extensively by lawyers), how the complainant was dressed, and drug or alcohol use by the complainant at the time of the incident. They might be based on stereotypical thinking about characteristics such as race, gender, and social class. The ban also encourages strong reliance on demeanor evidence, in another application of the law’s peculiar notion that, in the artificial atmosphere of the courtroom, truth can be determined by careful attention to cues such as hesitation, nervousness, shifty eyes, and other manneristic evidence about sincerity.\(^7\)

Character attributions are unavoidable. The question is whether accuracy would be promoted by admitting other-acts evidence, for example, evidence about whether the defendant sexually assaulted other women on dates, or whether the complainant falsely accused other men of rape. Merely pointing out the phenomenon of lay dispositionism does not provide an answer to the latter question. In fact, if the “interview illusion” analogy is valid, it would seem that history evidence would be a good supplement to the trial “interview.”\(^7\)

c. Problems Caused by the Trial Setting

Drawing inferences in the trial setting is in some ways more dif-

---


\(^7\) Consider also the literature on prediction of dangerousness. It is generally conceded that predictions are improved by consideration of “actuarial” information. See MONAHAN, supra note 10, at 71-72.

ficult than drawing inferences in ordinary life. Witnesses are re-
hearsed and dressed by lawyers. Experts are chosen tendentiously.
Evidence is generated with an eye toward litigation. Some parties
have more money than others for lawyers, experts, and investigators.
The impact of resource imbalance is hard to guess. In any event,
facts about resources are kept from the trier. Even the rules of evi-
dence can create mysterious and misleading holes. Lay jurors must
somehow account for the impact of a complicated and unfamiliar sys-
tem for investigating facts and forcing testimony.

These problems affect all testimony, not just other-crimes evi-
dence. For example, in an alibi defense case the prosecution may
have influenced eyewitnesses who identify the defendant in ways that
are hard to uncover or evaluate.

One feature of the trial setting does have a special impact on
other crimes evidence. In investigating unsolved crimes, the police
are likely to look for offenders who have prior records of similar
crimes. Unless the fact-finder takes this selectivity into account, the
evidence of prior crimes may be misleading. The task, however,
does not seem to be more difficult than other tasks routinely given
the trier, such as evaluating the accuracy of testimony of informers
who are in witness protection programs.

Some features of the trial setting may actually aid the evaluation
of character evidence. Cross-examination, discovery, and investiga-
tion may reveal context that would remain hidden in ordinary life.
Formal procedures for presenting two sides and reminding the trier
of evidentiary dangers may mitigate some of the harm that character
evidence might do in ordinary decision-making.

Some commentators have noted that the defendant, in deciding
whether to plead guilty, might take prior crimes evidence into ac-
count. If so, the fact that the defendant pled not guilty in the face of
other-crimes evidence may diminish the inferential value of the
similar offense evidence. This postulated effect has been advanced as
one reason for excluding evidence of other crimes. This insight,
while interesting and clever, does not really show how other crimes
evidence is uniquely dangerous or likely to mislead. Instead, it shows

78. "Unless the jury realizes that any criminal defendant is much more likely to have
committed prior bad acts than a randomly selected person from the general population, it
may overestimate the probative value of specific acts evidence." RONALD J. ALLEN ET
79. See RICHARD O. LEMPERT & STEPHEN A. SALTBURG, A MODERN APPROACH
TO EVIDENCE 217 (2nd ed. 1982); ALLEN ET AL., supra note 78, at 303. In fact, one
commentator has cautiously suggested that the absence of plea bargaining may make
similar offense evidence less dangerous in some Continental systems than it is in the
American system. See Mirjan Damaška, Propensity Evidence in Continental Legal Sys-
tems, 70 CHI.-KENT L. REV. 55, 60 (1994).
how hidden institutional pressures, of which plea bargaining is one, tend to complicate and confuse all common-sense inferences about guilt and innocence. The probative value of any incriminating evidence is diminished by the fact that in the face of such evidence, the defendant decided to risk the dangers of trial. For example, if the defendant pled not guilty in the face of DNA evidence, one might infer a greater than usual danger of error in the DNA testing procedure, on grounds that the defendant would not go to trial in the face of overwhelming evidence unless motivated by a sense of innocence.

At any rate, the postulated effect ignores the fact that other crimes evidence has an impact on both prosecution and defense. The prosecution and defense must both make evaluations of the strength of a case. The evaluations will be based on a combination of the probability of conviction and the probable severity of punishment. A record of past offenses increases the probable severity of punishment, hence giving the prosecution an extra bargaining chip. That will cause the prosecutor to bargain higher, sometimes causing a guilty defendant to decide to roll the dice rather than accepting the certainty of severe punishment. If a combination of overwhelming evidence of guilt and a record of prior crimes makes the prosecutor certain that she can win both a conviction and the desired sentence, the prosecution might refuse to bargain at all. That would make trial a free roll for the defendant, who would choose to plead not guilty regardless of consciousness of guilt.80

d. The Problematic Nature of A Priori Regulation of Inferential Error Prejudice

(i) Difficulty of Setting Rules Ahead of Time

One goal of the rule against character evidence is to prevent misdecision caused by cognitive mistakes or by the uncontrolled passions of the trier of fact. To believe that a rule excluding evidence prevents misdecision, one must combine pessimism about the ability

80. Professor Craig Callen has pointed to another danger connected with the trial setting. He has suggested that if character evidence is admitted, the jury may think that the judge is warranting that the evidence is probative and worth considering. See Craig R. Callen, Simpson, Fuhrman, Grice, and Character Evidence, 67 U. COLO. L. REV. 777, 781 (1996). That may be true, but the danger could be met with measures less drastic than exclusion, such as judicial comment, in cases in which the judge has a dim view of the evidence. In any event, while it is dangerous to admit, it is also dangerous to exclude. The jury may draw negative inferences from the absence of evidence. These inferences may be the product of misinformation, particularly where the existing rules permit bad acts of one party to be admitted but require that the bad acts of the opponent be excluded. See infra text accompanying note 115.
of humans to make accurate decisions in cases actually before them with a striking optimism about human foresight, one that endows lawmakers with the ability to create rules ahead of time that will improve decisions in cases that have not yet arisen. In a natural environment where new combinations of facts arise without warning and the trier must decide a case whatever its characteristics, any flat rule aimed at preventing mistaken inferences will collide with cases in which it cannot be applied flatly without bringing strange results.

General rules of law are often worth the price of hard results in special cases. But usually the reason for accepting hard results is that rules shape conduct and that rigidity aids planning. Rules aimed at preventing mistakes in a trier’s reasoning about the past are unique. When the purpose of a rule is to avoid mistaken findings about historical facts, and not to send a message about how to behave, then it is difficult for the rule’s administrators to feel that they owe it to the future to sacrifice present justice in order to give the rule clear boundaries. The result of trying to control reasoning about behavior with a priori rules will be a complex structure of rule and exception, hard to understand and hard to administer.  

(ii) The Questionable Assumption of Judicial Superiority

Another option, of course, is not to set rules ahead of time, but to highlight the area for scrutiny, telling the judge to screen out evidence that might lead to cognitive error. This approach has obvious limits in bench trials, where it would require judges to protect themselves from cognitive mistakes by being wiser in screening than they are in deciding cases on the merits. In jury trials, arguably the judge is better able than the jury to see the flaws of the testimony. But while judges may be experts about the adversary system (and hence able to screen to prevent or counter adversarial abuse) there is no reason to think that they are experts in avoiding attribution error in evaluating character evidence. Appointment by the governor does not convey immunity to flaws in reasoning about behavior. Even if judges were less prone to inferential error than juries, giving them the discretionary power to make or break a case by controlling the evidence would raise other dangers. The institution of the jury trial is designed for a world where venality and oppression are dangers as great as mistake. By using an ad hoc body chosen largely by the forces of chance, we reduce the dangers of favoritism, political pres-

81. Of course, discretionary rules to be applied in limine can have similar problems, to the extent that they require the judge to make adjustments before hearing all the evidence; but at least the rules are being applied in light of known facts, with the judge retaining the authority to change rulings.
sure, and corruption. Rules aimed at preventing inferential error can undermine the usefulness of jury trial in guarding against these other dangers. The learning about prediction of dangerousness by mental health experts suggests that beating the lay person at the job is a daunting task even for those trained in it. Judges are relatively isolated and nondiverse; a diverse body of lay fact-finders might well be superior at assessing questions such as whether a defendant’s sale of LSD while in high school is worth considering in deciding whether the defendant sold heroin five years later. An assessment of dangers of surprise and adversarial abuse requires knowledge of the legal system; an assessment of consistency of behavior requires knowledge of how the rest of the world works.

Jurists have overemphasized the danger of inferential error in excluding character evidence. Undoubtedly the evidence does create that danger in some cases. The literature on attribution error gives some cause for concern, in the trial setting and elsewhere, that character evidence will encourage overconfidence in character attributions. At trial, character evidence may also lead to distraction from the hard task of evaluating circumstantial evidence. For example, in a case in which challenged forensic science evidence connects the defendant with a crime scene, the trier of fact might become more interested in the story of the defendant’s life and character, and less willing to do the hard mental work of understanding the forensic science. But were inferential error the only danger, the difficulty of trying to guard against it by rule and the danger of allowing discretionary exclusion would not justify excluding other-crimes evidence.

(2) Nullification Prejudice

The rules against character evidence make more sense if one considers them to be aimed at preventing nullification of the substantive law than if one sees them as directed against inferential error. If character could be explored freely, triers would be tempted to give litigants what they deserve, not what the law requires.

A focus on particular transactions and occurrences, rather than upon one’s lifetime merit, helps lawmakers send messages about primary conduct that can be understood and obeyed. It may even enhance the economic efficiency of legal precedent, since the focus on parties as the representatives of those engaging in an activity (rather than a focus on personal worth) encourages law-making judges to consider the economic value of competing activities.


83. See RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 496 (3d ed. 1986).
In an earlier article, I wrote that it made sense to give judges a role in preventing nullification, since their identification with the goals of the law makes them less vulnerable to nullification prejudice than juries. One can question this point by saying that the character ban applies alike to judge and jury trials, hence calling into question whether it is based on suppositions about a judge's superior resistance to nullification prejudice. But the rules apply with more force in jury trials than in nonjury trials for the obvious reason that in nonjury trials, the judges must hear the evidence while screening it. If lawmakers were strongly concerned about the impact of character evidence upon judges, they would have to go much further than they have now—for example, by bifurcation of proceedings to separate the evidence-screener from the trier of fact, or at least by eliminating the generous presumption that a judge who admitted inadmissible evidence subsequently disregarded it in reaching the verdict.

In any event, even if judicial nullification is as great a danger as jury nullification, the character evidence rules would still have some role to play in preventing nullification. Even a judge eager to nullify would feel some limits on exploring the grounds for nullifying. She could not, for example, by bifurcation of proceedings to separate the evidence-screener from the trier of fact, or at least by eliminating the generous presumption that a judge who admitted inadmissible evidence subsequently disregarded it in reaching the verdict.

B. Cost-effectiveness

Evidence of general propensity takes a long time to develop (if done carefully) and is of limited probative value when used to show conduct in dissimilar situations. Concern for time explains many of the otherwise peculiar elements of the current rule. It may also explain why courts are considerably more permissive in situations in which prejudice is a problem but cost is not, for example, when dealing with the question of whether two offenses by the same person may be joined or whether to permit an advocate to make insulting

84. See Park, supra note 14, at 771.
85. On the "generous presumption," see, for example, United States v. Impson, 562 F.2d 970, 971 (5th Cir. 1977) ("A judge, sitting as a trier of fact, is presumed to have rested his verdict only on the admissible evidence before him and to have disregarded that which is inadmissible"). But cf. Anderson v. Smith, 751 F.2d 96, 106 (2d Cir. 1984) (reversal on grounds that judge in bench trial might have considered evidence erroneously admitted in violation of the Constitution).
86. These features include the preference for reputation/opinion testimony over specific acts, see FED. R. EVID. 405(a), and rules allowing witnesses to be questioned but not contradicted, see FED. R. EVID. 608(b).
87. The federal provision on joinder of charges against a single defendant is broadly permissive. FED. R. CRIM. P. §8(a) provides:
Two or more offenses may be charged in the same indictment or information in
comments about a party's character during final argument. 88

In the absence of time and money constraints, we could go into the defendant's whole life in depth, even allowing competing expert testimony offering clinical assessments of an individual's character. 89 But given limited resources, there is good reason to hesitate before opening the door to expensive, adversarially-generated evidence of dubious probative value. 90

One might think that, in the absence of any danger of prejudice, the problem of cost would take care of itself, since the self-interest of adversaries would prevent them from offering expensive but unconvincing evidence. But society's calculus may differ from that of an adversary, especially in a system where the state bears many of the expenses of litigation. 91 A wealthy litigant whose only hope was

a separate count for each offense if the offenses charged, whether felonies or misdemeanors or both, are of the same or similar character or are based on the same act or transaction or on two or more acts or transactions connected together or constituting parts of a common scheme or plan.

Were the charging rule to be transferred to the law of character evidence, one would see other-crimes evidence admitted when the other crime was of the "same or similar character" as the crime charged—an opening far greater than the one created by Rules 413-415. For cases showing willingness to join, see United States v. Coleman, 22 F.3d 126, 131-35 (7th Cir. 1994)(four separate felon-in-possession cases could be joined together because of similarity, despite lack of evidentiary overlap); United States v. Jines, 536 F.2d 1255, 1257 (8th Cir. 1976) (holding joinder of possession of unregistered firearm and distribution of heroin charges proper; acts were committed same day and observed by same two detectives); United States v. Pietras, 501 F.2d 182, 185 (8th Cir. 1974) (holding that various bank robbery charges properly joined with possession of unregistered firearm which was not used in the robbery because gun was found in van used in robbery); Edwards v. Squier, 178 F.2d 758, 759 (9th Cir. 1949) (holding interstate transport of a stolen vehicle and stolen securities, while not in the same transaction, nonetheless properly joined because both involve interstate transport); United States v. McCoy, 848 F.2d 743, 744-45 (6th Cir. 1988) (holding that joinder of armed bank robbery and subsequent unarmed bank robbery proper); United States v. Bourassa, 411 F.2d 69, 74-75 (10th Cir. 1969) (holding that possession and passing counterfeit coins and bail jumping properly joined, or at least harmless error when bail jumping arose from failure to appear on counterfeit charge). But see United States v. Hubbard, 61 F.3d 1261, 1269-71 (7th Cir. 1995) (holding that distribution and conspiring to distribute cocaine and possession of a firearm by a felon improperly joined because the firearm was discovered 17 months after the narcotics transaction).

90. The concern is nothing new; Hamlet, who faced problems seemingly remote from this one, nonetheless enumerated "the law's delay" as one reason why suicide was so tempting. See WILLIAM SHAKESPEARE, HAMLET, act 3, sc. 1, at 886 (W.J. Craig ed., Oxford Univ. Press 1969).
character evidence might have a marginal utility function different from that of the judge, jury, and society. But a stronger answer, explored in the next section, is that considerations other than accuracy may influence the litigant.

C. Preventing Adversarial Misconduct and Distortion

Adversaries have temptations to use and misuse character evidence that do not depend upon any suppositions about the cognitive flaws of the fact-finder.

(1) Bullying, Threats, Revenge

Character evidence is usually offered to support character attack. Being the victim of character attack can be unpleasant regardless of its effect on the fact-finder. Few litigants or witnesses would relish a far-ranging examination of character at the hands of an adversary, particularly at the hands of a lawyer paid by a personal enemy. Character attack—even unappealing character attack that would not tempt the passions of the trier of fact—can be used to bully a party or witness. It can obstruct the bringing or the support of a meritorious claim. This danger has been recognized most expressly in the area of character protection known as “rape shield,” but of course it exists in other situations. Moreover, even where instrumental value is lacking, a party, and an overly compliant lawyer, might still engage in character attack simply to cause pain to an enemy.

(2) Surprise, Ambush, and Attendant Hazards

The danger of surprise, ambush, and attendant difficulties in meeting evidence sprung by an adversary are undoubtedly another reason for the character ban. To Wigmore, they were the “chief reason” why specific acts evidence was not allowed. One cannot de-

(1988)(refuting argument that self-interest of parties will cause them to produce the best available evidence).


To the general rule allowing the use of all circumstantial facts without giving prior notice, and refusing to recognize unfair surprise as a ground for the exclusion of evidence, there seems to be but one generally recognized exception, at common law. There is, as already noticed ($1847 supra), a special and palpable danger of undetectable fraud in allowing the moral character of an opposing party or witness to be evidenced by particular acts of misconduct, or particular falsities, when attempted to be proved otherwise than by cross-examination of
fend one’s whole life on the spur of the moment. Modern criminal
discovery and notice requirements mitigate but do not completely
solve the surprise problem. Requiring detailed notice far enough in
advance to permit full preparation would either be so burdensome
that it would often have the same effect as a rule excluding other-
crimes evidence, or it would interfere with other goals, such as the
goal of having a speedy and inexpensive trial.

(3) “Round up the Usual Suspects”

In criminal cases, the character evidence ban may also encourage
police and other investigators to allocate resources to something
more productive than “rounding up the usual suspects” or digging up
the “dirty laundry” of an accused. I do not put too much emphasis
upon this point, because I think that the character evidence rule is an
awkward way of directing the allocation of resources, and that it is
not always bad to scrutinize prior offenders or to look for similar past
conduct. But it is possible that the rule against character evidence
has an incidental benefit of encouraging the use and development of
incident-specific evidence, such as fingerprints and DNA profiles,
and of encouraging the parties to do the hard work of exploring that
evidence instead of digging up material for narratives about charac-
ter.

D. Goals Outside the “Rationalist Tradition”

In discussing the goals of the character ban, my perspective has
been consistent with what William Twining has called the Rationalist
Tradition of Evidence Scholarship. The rationalist model of adjudic-
ation, in Twining’s account, assumes that the prescriptive goal of
procedural law is “rectitude of decision”—that is, correct application
of the substantive law through accurate determination of past facts,
after careful and rational weighing of evidence by impartial decision-
makers. The descriptive assumptions of (some) adherents to this
tradition are that “[g]enerally speaking this objective is largely

---

the party or witness himself or by record of conviction for crime. Other reasons
of policy, however, combined to oppose such evidence; and accordingly it was
not merely prohibited unless after notice given, but prohibited unconditionally.

Id.

94. From the point of view of offenders, a system under which prior offenders were
released more quickly but subjected to greater scrutiny and suspicion after release might
be more merciful than the one we have now.

95. See WILLIAM TWINING, RETHINKING EVIDENCE: EXPLORATORY ESSAYS 32-91

96. Id. at 73 tbl.1.
achieved in a consistent, fair, and predictable manner.” This model assumes that accurate knowledge of the past is possible, that establishing the truth about the past is a necessary condition for achieving justice in adjudication, and that judgments about probable truth can be made by inductive reasoning. Twining states that the model also includes the assumption that “[t]he pursuit of truth (i.e., seeking to maximize accuracy in fact-determination) is to be given a high, but not necessarily an overriding, priority in relation to other values, such as the security of the state, the protection of family relationships or the curbing of coercive methods of interrogation.”

All of the rationales that I have described so far are consistent with the hypothesis that evidence law should aim at historical accuracy. The conflict between accuracy in litigation and other goals is sometimes an illusory one. Limits on character evidence aimed at witness protection (e.g., the rape shield rules) are consistent with ex ante pursuit of judicial accuracy. They encourage witnesses to come forward and participate. Limits that are aimed at equalizing adversarial opportunities can also be consistent with accuracy, because (1) the adversarial system is itself arguably an accuracy-serving system considering imperfections in human nature, such as a tendency to prejudgment and selective pursuit of information, and (2) equalization within the system can be seen merely as a way to make sure both sides have the chance to present information in the pursuit of accuracy.

In a broader sense, procedural rules aimed at cost-reduction also serve the cause of accuracy. If the pursuit of accuracy means establishing a regime of fact-finding that accurately determines whether substantive rules have been violated, then the decision to keep down the cost of litigation in a particular case is no more at war with a desire for accurate results than is a scientist’s decision to divide funds between two experiments rather than spend all on one. At some point along the cost spectrum, accurate enforcement of substantive law will be harmed by single-case perfectionism, because cost will drive cases out of the fact-finding system completely. Moreover, high time-cost harms the quality of the fact-finder, for example by making citizens flee the prison of jury service.

Of course, the conflict between accuracy and other goals is not always illusory. Commentators have sometimes attributed goals to evidence law that are hard to reconcile with a truth-seeking system, at least not without some very speculative consideration of indirect effects. For example, Professor Nesson has suggested that accept-

97. Id.
98. Id.
ability is a goal that rivals accuracy, and Professors Leonard and Taslitz have postulated that "catharsis" is a legitimate goal.

(1) Departures from the Rationalist Tradition—as Description

As a descriptive matter, there is nothing particularly startling about the idea that evidence rulemakers and rule enforcers sometimes pursue goals other than accuracy. They surely do, just as surely as politicians sometimes sacrifice principle for expediency. The question is one of degree and frequency. Nesson’s Acceptability article states a strong version. Nesson sees the judicial system as often preferring the pursuit of acceptable verdicts over the pursuit of truth. He thinks that this preference is displayed in the very text and structure of announced rules (as opposed to only being exemplified by sub rosa misapplications of rules). Nesson offers the following examples to show that the pursuit of acceptability leads to different results than the pursuit of truth: the rules preventing direction of a verdict when a case turns on credibility, the hearsay rule, the conjunction problem, the attorney-client privilege in criminal cases, and the law’s uneasiness with reaching verdicts based on “naked” statistical evidence. All of these examples are susceptible to plausible (or better) explanations based upon a truth-finding rationale.

A weaker version of Nesson’s descriptive thesis is undoubtedly true, even truistic. Of course human actors, including judges, give weight to acceptability in planning conduct, and they sometimes bend

100. See Leonard, supra note 28, at 2-3.
101. See Taslitz, supra note 89, at 60-63.
102. See Nesson, supra note 99.
103. Nesson’s “acceptable” verdicts are verdicts satisfactory to the public at large, as believable (though inaccurate) statements of historical fact:

Many of the procedures of our legal system are best understood as ways to promote public acceptance of verdicts. These procedures facilitate both the initial and the continued acceptance of the verdict as a statement about what actually happened. Judges and commentators alike tend to underestimate the role that acceptance of verdicts may play in accounting for certain procedures, preferring instead to rationalize evidentiary and procedural rules as means to advance the search for truth. Many rules are indeed explicable in terms of a truth-seeking rationale [for example, the rule of relevance]. But, on close inspection, some procedures that are rationalized as truth-seeking devices are better understood as means to promote public acceptance of verdicts.

Nesson, supra note 99, at 1368-69 (footnotes omitted).

the truth in pursuit of popular approval. But that does not mean that there is a systemic and serious bias against accuracy.

First, accuracy and acceptability are not often at war. One could plausibly argue that desire for acceptability creates a bias toward accuracy in our trial system—a bias that makes verdicts more accurate than permitted by law. For example, it is possible that desire for substantive accuracy has trumped procedural law in cases enforcing rules that purport to exclude illegally obtained evidence.

Second, it seems to me that serious inroads on accuracy would be so harmful to the interests of the powerful as to evoke their heated resistance. To say that judges pursue accuracy is only to say that they seek to carry out the mandates given by substantive law. A system that in general enforces the substantive law—that determines what happened in order to uphold governmental mandates telling citizens and subjects how to conduct their primary affairs—is one that in general serves the interests of those who hold power.

Of course, the evidence rules are far from perfect, and cynics can help us explain the reason for some of the imperfections. The cynical view that makes the most sense to me is that aberrations in the pursuit of truth are caused by the self-interest of in-house rulemakers (judges and lawyers) when they are not acting as faithful agents of other holders of power. The self-interest of "Judge & Co." may account for some defects in our system, such as rules that favor lawyers at the expense of witnesses, that expand the attorney-client privilege, or that permit neglect of such obvious truth-enhancing possibilities as the use of neutral experts. The harm of such rules is spread broadly, or at least the people harmed cannot see it coming, and the beneficiaries are a small group of influential repeat players. But with that caveat, I find it hard to detect a systematic bias against accuracy in evidence rules.

In the character context, one possible example of codified bias against accuracy is the "mercy rule" allowing criminal defendants to put in evidence of good character in situations in which the prosecution may not. It seems entirely plausible to me, however, to suppose that rule to be consistent with the pursuit of accuracy, on grounds that evidence of good character does not have the same power to in-

105. In the context of expert witnesses, the failure to embrace neutral experts may be partly due to the desire of lawyers to control their witnesses and the reluctance of judges to undertake the work (early in a case that may settle) of appointing appropriate experts. Cf. Samuel R. Gross, Expert Evidence, 1991 Wis. L. Rev. 1113, 1221.

106. Cf. William N. Eskridge & Phillip P. Frickey, Legislation: Statutes and the Creation of Public Policy 611 (1st ed. 1988) (when there are concentrated benefits and distributed costs, one can expect organized activity supporting legislation, little organized opposition, and a response of either self-regulation or distribution to the organized group).
vite nullification prejudice as evidence of bad character. In addition, the evidence is likely to be reliable because in offering it the defendant lays open her entire life to examination, including questioning of her character witness(es) about specific acts of conduct. Because of the power of the prosecution's weapons of counterattack, lawyers will be reluctant to use the evidence except where the prosecution is known to lack any evidence that the defendant committed other bad acts like those with which she is now charged. As such, the evidence of good character is, functionally, really a stand-in for evidence of a lifelong pattern of not committing the type of act charged. This is a pattern whose frequency and stability might satisfy the requirements of even those who are skeptical about the value of cross-situational character evidence based on single instances.

If that explanation for the "mercy rule" does not satisfy, then perhaps one does need to postulate a goal other than pursuit of accuracy. The alternative that seems most appealing to me is not the concept of catharsis, but a more cynical idea: rulemakers tend to be respectable and well-connected, and hence have a bias toward sympathizing with the sort of respectable and well-connected defendant who is likely to benefit from the rule.

(2) Departures from the Rationalist Tradition—as Prescription

Professor Nesson's Acceptability article could also be read as a dissent from the prescriptive dimension of the rationalist tradition. His ultimate conclusion, though stated in qualified, depersonalized, and contingent language, seems to be that truth-seeking should sometimes yield to the pursuit of publicly acceptable verdicts. This proposition is problematic. The public does not wish for its courts to act as self-serving public relations agencies at the expense of finding the truth. If one believes that acceptability should trump accuracy, then one must give jurists a hidden agenda, telling them to pursue se-

107. One who is absolutely committed to the process of ascertaining and testing the truth, and who would thus shun any concession of the search for truth to the production of acceptable verdicts, may find that he does so at the expense of other important values. He may discover that extremes in the pursuit of truth can impair the system's capacity to generate acceptable verdicts and thus undercut its ability to project the norms embodied in the substantive law. The discomfiting thought that our quest for the truth must not weaken our drive toward acceptable verdicts undermines the comfortable position that our drive toward acceptable verdicts should not compromise our quest for the truth.

Nesson, supra note 99, at 1392.

108. For a stronger statement by a commentator who appears to embrace the concept of trial as theater with glee, see Kenneth W. Graham, Jr., "There'll Always Be an England": The Instrumental Ideology of Evidence, 85 Mich. L. Rev. 1204 (1987).
cret goals through devious means. Given that mandate, they might end up creating a system that served their own interests more than it served the public's. At any rate, they might get caught at the trick, thus defeating the goal of acceptability. A milder form of the acceptability thesis can, however, serve to remind us that we should not get too far ahead of common sense, that public satisfaction with verdicts is a blessing, and that pursuing solutions appealing only to an intellectual elite may have adverse side effects.

Returning to the specific topic of character evidence, it seems to me dangerous to justify wrinkles in the law of character evidence on the idea that "catharsis" should sometimes trump truth-seeking. A postulated need for "catharsis" could be used by power holders to justify any decision-making process, from gladiators to divine inspiration. The quest for accuracy, even if doomed to failure in many cases and debatable success in others, puts limits on what rulemakers and rule enforcers can do, asking them to persuade us that their chosen process is consistent with a rational attempt to achieve accurate verdicts.

III. Particular Problems with the Current Rules

A. Complexity

Every evidence teacher is familiar with the horrifying complexity of the character evidence rules. The system of exceptions and exemptions has been made even more convoluted by the advent of a whole new class of exceptions for sex crime cases.

Complexity may be inherent in any attempt to regulate reasoning about behavior with firm *a priori* rules. Evidence of prior bad acts, for example, can link up with other evidence in an infinite number of ways. The appearance of unexpected constellations of facts, and the complexity of the process of drawing inferences in a natural environment, lead to a complicated structure of rule and exception, especially since judges (rightly so) will have difficulty seeing the benefits of adhering to a flat, simple, rigid rule.

B. Evasions, Misuse, and Mystification

The complexity and confusion attending the character ban are enhanced by numerous evasions, even in published appellate cases, of the supposed ban on general propensity evidence. In applying Rule 404(b) and its state analogues, some judges act as if a "plan" exists

whenever a defendant has done bad things twice, or that bad intent on one occasion is always admissible to show bad intent on another, even if the bad intent is inferred from general propensity to commit a type of crime. The rule against character evidence leads to other problems. It encourages unintelligible limiting instructions, for example, to use the evidence not to show character, but to show intent or plan. It seems to evoke ritual incantations of lists of permissible purposes, rather than meaningful on-the-record weighing of considerations of cost and prejudice against the factors that psychology and common sense tell us are important—such as similarity of situation, frequency of conduct, and proximity in time. Sometimes these are considered—it is hard to be consistent in resisting common sense—but the formula in 404(b) does not help put them in the forefront and sometimes seems to do more harm than good.

C. Inconsistent Treatment of Like Cases

The rules create inconsistencies in the treatment of like situations. Character impeachment of witnesses is one example—the rules allow a conviction for a crime of violence to be used to attack character for truthfulness but not character for violence. Another


111. Cases involving drug dealers provide the most dramatic examples of using cross-situational propensity evidence to show intent. See United States v. Broussard, 80 F.3d 1025, 1039-40 (5th Cir. 1996) (no abuse of discretion to admit prior drug offenses, despite remoteness in time as to one defendant and lack of similarity as to another); United States v. Bermea, 30 F.3d 1539, 1562 (5th Cir. 1994), cert. denied, 115 S. Ct. 1113 (1995) (noting general receptiveness toward drug offenses as 404(b) evidence); United States v. Perkins, 94 F.3d 429, 433-34 (8th Cir. 1996), cert. denied, 117 S. Ct. 1004 (1997) (prior contact of defendant with crack cocaine admissible to show his knowledge and intent); United States v. Moore, 98 F.3d 347, 350 (8th Cir. 1996) (Loken, J.) (a "mere presence" defense to a drug charge puts intent in issue, so other drug offenses are admissible to show intent); United States v. Butler, 102 F.3d 1191, 1195-96 (11th Cir. 1997), cert. denied, 117 S. Ct. 1712 (1997) (held, in an opinion citing cases on both sides, evidence of personal drug use more than three years before an unrelated prosecution for distribution conspiracy is admissible to prove intent). But see United States v. Beasley, 809 F.2d 1273, 1279-81 (7th Cir. 1987) (Easterbrook, J.) (one drug offense not always admissible to prove another, even when intent is in issue; remanded for on-the-record 404(b) balancing).

112. See, for example, Judge Ito's limiting instruction in the O.J. Simpson case, telling the jury that the evidence of spousal abuse could not be "considered by you to prove that the defendant is a person of bad character or that he has a disposition to commit crimes," but that it was admissible only for a laundry list of other purposes, such as showing plan, scheme, motive, intent, or identity. 1995 WL 672668, at *12 (Cal. Super. Ct. L.A. County, Sept. 22, 1995).

113. For commentary on this anomaly, see Friedman, supra note 34, at 637-39; Office
is the inconsistency between the treatment of sexual assault and non-
sexual violence introduced by Rule 413.  

D. Perverse Side Effects

The rules also have perverse side effects. For example, under the current rules the defendant will sometimes choose not to testify because of fear that testimony will lead to impeachment with prior convictions. Were prior crimes admissible for substantive purposes, this side effect would disappear. The existing rules may also lead to mistaken jury inferences based on the absence of evidence. For example, a prosecution witness might be impeached with prior convictions in a case in which prior crimes of the defendant did not come to light because the defense chose not to put him on the stand. The jury might make character attributions about both defendant and witness, based on incomplete information. Alternatively, it is possible that having heard of exclusionary rules, jurors might even draw mistaken inferences from mistaken suppositions about their effects—assuming, for example, that prior crimes would never be admissible.

In short, there are a host of particular problems with the existing rule structure. The problems—except complexity—might be dealt with by incremental reform. The growth of the grotesque structure may, however, be a sign that conventional rationales are unconvincing, or that the project of a priori inference regulation is failing.

IV. Cures and Palliatives

A. A Crime-by-Crime Approach to Relaxation of the Ban

(1) Rationale of a Crime-by-Crime Approach

A "crime-by-crime" approach is one that experiments with changing the ban on character evidence crime by crime. Using a crime-by-crime approach means letting in prior crimes of the same general nature as those charged in the present case, and then allow-
ing propensity reasoning based on those prior crimes.

It is doubtful we would be taking such an approach seriously but for the model Congress has recently given us in Federal Rules of Evidence 413 and 414. Those rules allow prior rapes to be used for character reasoning against a defendant charged with rape, and prior instances of child molestation to be used against a defendant charged with child molestation. Rule 403 is still available for fall-back arguments to exclude crimes that stem from quite different circumstances or motives, or that are remote in time, but even so the new rules represent a significant departure from prior law.

The new rules also create anomalies. For example, if a defendant is accused of rape and murder, the defendant’s prior rapes are admissible for character purposes but not his prior murders. A crime-by-crime approach could be used to smooth out this sort of inconsistency, either by exempting additional crimes from the character ban or by cutting back on the scope of those now exempted.

The crime-by-crime approach has some attractive features aside from promoting incremental consistency. The criminal code’s classifications establish a hierarchy of seriousness and identify crimes whose harmful nature justifies moral condemnation, labeling, and extended punishment. If a crime is serious enough to justify spending enormous state resources detaining the accused for a long period of time once he has been convicted, perhaps it is serious enough to justify the time and effort needed to gather and examine specific-act character evidence in its situational and psychological context. Moreover, quantitative data about recidivism, incidence, and victimization is collected systematically according to crime categories. We might be able to take advantage of that information in determining which crimes are the best candidates for a relaxation of the ban


118. I would not call it odd, however, to admit prior rapes in a consent-defense case even if evidence of other crimes is excluded when offered for character purposes. See infra text accompanying notes 134-35.
against character evidence.

(2) Guidelines for a Crime-by-Crime Approach

In formulating a crime-by-crime approach, lawmakers might consider the seriousness of the crime, the need for the evidence, and the comparative propensity of persons who commit that type of crime.

a. Comparative Propensity

(i) Recidivism and Comparative Propensity

Recidivism studies measure the rate of reversion to criminal behavior after a convicted perpetrator has been released from custody.119 Commentators have used information about recidivism rates as one yardstick for measuring the probative value of different types of character evidence. They have made the understandable assumption that the higher the rate of recidivism for a crime, the better the case for admission of propensity evidence when a defendant has been charged with that crime.120 For example, Professor Edward Imwinkelried asks us to think of the 1994 “crime bill” as an experiment in selective abolition of the rule against character evidence.121 He


120. See Thomas J. Reed, Reading Gaol Revisited: Admission of Uncharged Misconduct Evidence in Sex Offender Cases, 21 AM. J. CRIM. L. 127 (1993). Reed devotes several pages to recidivism data and, although it is not always clear how his information about recidivism relates to his other conclusions, in general he appears to believe that the higher the rate of recidivism after conviction for a particular crime, the better the case for admitting specific act evidence about that crime, and that the case for admitting sex offense evidence is bolstered by studies suggesting a high recidivism rate. See id. at 146-56. Compare James M. H. Gregg, Other Acts of Sexual Misbehavior and Perversion as Evidence in Prosecutions for Sexual Offenses, 6 ARIZ. L. REV. 212, 233 (1965), which cites recidivism studies indicating a low rate for sex crimes in support of the position that there should be no special exception for sex crimes cases. Cf. James S. Liebman, Symposium, Violent Crime Control and Law Enforcement Act of 1994: Proposed Evidence Rules 413 to 415—Some Problems and Recommendations, 20 U. Dayton L. REV. 753, 756 (1995)(noting that despite a reportedly high recidivism rate for burglars and robbers, evidence of these types of crimes is excluded, whereas evidence of previous sexual assault crimes is admissible under Federal Rule of Evidence 413); Bryden & Park, supra note 115, at 572-73 (noting that recidivism studies do not affirmatively support distinction between sex crimes and other crimes); see also Commonwealth v. Boulden, 116 A.2d 867, 874 (P.A. Super. Ct. 1955) (citing recidivism data in support of position that “there is no more reason to admit prior offenses to show depravity or propensity in a sex case than in any other case.”). But see Lannan v. State, 600 N.E.2d 1334, 1336-37 (Ind. 1992) (noting recidivism data but giving it little weight).

121. Edward J. Imwinkelried, Undertaking the Task of Reforming the American Character Evidence Prohibition: The Importance of Getting the Experiment Off On the Right
criticizes Congress for selecting the wrong crimes for the experiment. He argues that sex crimes have "minimal probative value as predictors of the accused's conduct," and suggests that "[i]t would make far more sense to initiate the experiment by selecting crimes with higher recidivism rates." Noting that in a leading study the reported recidivism rate for burglary was 31.9%, while for rape it was 7.7%, he suggests that burglary would have been a better candidate than rape. Similarly, Professor Katharine K. Baker states that:

Advocates of Rule 413 also unabashedly and without proof suggest that rapists are more likely than other criminals to repeat their acts. The evidence that we have is to the contrary. A 1989 Bureau of Justice Statistics recidivism study found that only 7.7% of released rapists were re-arrested for rape. In contrast, 33.5% of released larcenists were re-arrested for larceny; 31.9% of released burglars were re-arrested for burglary; and 24.8% of drug offenders were re-arrested for drug offenses. Only homicide had a lower recidivism rate than rape. It is true that released rapists are more likely than other released prisoners to be re-arrested for rape, but that rapists are more likely than others to rape again does not distinguish rapists from other criminals. Larcenists are twenty-five percent more likely to be re-arrested for larceny than rapists are to be re-arrested for rape. Arguing from the statistics, a crime-based prior act exception is better suited to larcenists and drug offenders than to rapists.

Recidivism data can indeed be helpful in assessing the probative value of other-crimes evidence. However, an assessment of the probative value of other-crime character evidence requires a comparison of the criminal propensity of prior offenders with the criminal propensity of other persons. To estimate comparative propensity, one needs to consider more than naked recidivism data. When a given crime has a low incidence in the general population, the probative value of evidence of another instance of the same crime will be greater than would have been the case had the crime been more

122. See id. at 297.
123. Id. at 301.
124. The principal study on which Professor Imwinkelried relied was a 1989 Bureau of Justice Statistics study. See Id. (citing BECK STUDY, supra note 11, at 6). The study tracked 100,000 released prisoners for three years, using rearrest rates as a measure of recidivism. I should note that Professor Imwinkelried did not base his whole thesis on recidivism. For example, he thought that sex crime evidence was more likely to create jury prejudice than evidence of prior burglaries. See id. at 296-97.
125. Katharine K. Baker, Once a Rapist? Motivational Evidence and Relevancy in Rape Law, 110 HARV. L. REV. 563, 578-79 (1997) (footnotes omitted); cf. Bryden & Park, supra note 115, at 572-73 (advancing a similar point, but not carrying it as far; the present author's thinking has changed upon further consideration of the implications of recidivism data).
common, even if the recidivism rate for the crime is low.

As an illustration, suppose burglary by parachute to be a crime separately defined by the criminal code. Were that the case, a prior instance of burglary by parachute, one that had no role in focusing suspicion on the defendant charged again with the same crime, would be more probative in a burglary by parachute case than a prior instance of common burglary in a common burglary case. That would be so even if burglars by parachute had the same absolute propensity to repeat the crime of burglary by parachute that common burglars had to repeat the crime of common burglary. That is, if 30% of all burglars by parachute repeated the crime of burglary by parachute, and 30% of all common burglars repeated the crime of common burglary, the other crime evidence would still be more probative in the burglary by parachute case. The absolute propensity of the parachute burglar would be the same as that of the common burglar, but the parachute burglar’s comparative propensity would be much higher. A person with a burglary by parachute record would be many times more likely to commit that crime than a person chosen at random from the general population (or from the population of all burglars). This common-sense induction is reflected in the accepted rule that crimes that have a distinctive modus operandi are admissible to show propensity.126

Consider Table 1, a table from the recidivism study cited by Professors Imwinkelried and Baker.127

126. I am here using “propensity” simply to refer to a proclivity or tendency of the defendant. In modus operandi cases, like habit cases, evidence is offered to show a propensity that is considered to be more narrow than a trait of character. For example, evidence of a propensity to drown spouses in the bathtub, see Rex v. Smith, 11 Cr. App. R. 229, 84 L.J.K.B. 2153 (1915), or to warm freon by heating it with a coil, see Halloran v. Virginia Chemicals Inc., 361 N.E.2d 991, 996 (N.Y. 1977), is considered not to be character evidence. American courts would, however, avoid the use of the word “propensity” in describing evidence considered to be admissible under the modus or habit concepts, because “propensity” is considered a synonym for “character,” and if deemed evidence of “character” the evidence would be inadmissible.

127. Table 1 is a reproduction of Table 9 from the Beck Study, see BECK STUDY, supra note 11, at 5. The first footnote to the table has been modified (as indicated in the footnote).
Table 1. Rearrest Rates of State Prisoners Released in 1993

<table>
<thead>
<tr>
<th>Rearest charge</th>
<th>Total all offenses</th>
<th>Violent offense</th>
<th>Property offense</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Murder</td>
<td>Rape</td>
</tr>
<tr>
<td>All charges</td>
<td>62.5%</td>
<td>59.6%</td>
<td>42.1%</td>
</tr>
<tr>
<td>Violent offenses</td>
<td>22.7%</td>
<td>30.4%</td>
<td>21.6%</td>
</tr>
<tr>
<td>Homicideb</td>
<td>1.6</td>
<td>2.8</td>
<td>6.6</td>
</tr>
<tr>
<td>Rape</td>
<td>.9</td>
<td>1.7</td>
<td>.8</td>
</tr>
<tr>
<td>Robbery</td>
<td>9.9</td>
<td>14.1</td>
<td>7.0</td>
</tr>
<tr>
<td>Assault</td>
<td>12.6</td>
<td>15.7</td>
<td>10.5</td>
</tr>
<tr>
<td>Property offenses</td>
<td>39.7%</td>
<td>32.1%</td>
<td>16.8%</td>
</tr>
<tr>
<td>Burglary</td>
<td>18.4</td>
<td>12.6</td>
<td>6.4</td>
</tr>
<tr>
<td>Larceny/theft</td>
<td>21.2</td>
<td>16.3</td>
<td>7.4</td>
</tr>
<tr>
<td>Motor vehicle theft</td>
<td>5.5</td>
<td>4.0</td>
<td>2.5</td>
</tr>
<tr>
<td>Fraud</td>
<td>6.5</td>
<td>4.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Drug offenses</td>
<td>16.6%</td>
<td>14.8%</td>
<td>9.1%</td>
</tr>
<tr>
<td>Public-order offenses</td>
<td>29.9%</td>
<td>29.0%</td>
<td>19.2%</td>
</tr>
<tr>
<td>Number of released prisoners</td>
<td>106,216</td>
<td>56,769</td>
<td>3,258</td>
</tr>
</tbody>
</table>

Note: The numerator for each percent is the number of persons rearrested for a new charge, and the denominator is the number released for each type of offense. Detail may not add to totals because persons may be rearrested for more than one type of charge.

a. Includes [non]negligent manslaughter. (The bracketed prefix corrects a clerical error in the originally published table. The footnote in the originally published table stated that “murder” included “negligent manslaughter.” It should instead have stated that “murder” included “nonnegligent manslaughter.” Telephone Communication from Allen J. Beck to Roger C. Park (Oct. 9, 1997)).

b. Includes murder, nonnegligent manslaughter, and negligent manslaughter.
Examining this table, note the following reported same-crime recidivism rates:

- Rape: 7.7%
- Burglary: 31.9%
- Larceny: 33.5%
- Drugs: 24.8%

From these figures, both Professors Imwinkelried and Baker separately conclude that rape is a relatively unpromising candidate for a same-crime exception because of rape's relatively low same-crime recidivism rate.

One problem with this reliance on naked recidivism data can be seen by examining the other figures in the table. A higher percentage of released rapists were re-arrested for burglary (12.7%) than for rape (7.7%). That cross-crime recidivism figure does not mean that evidence that an accused had previously committed rape would be more probative of guilt in a case charging burglary than in a case charging rape.

If one had to make decisions about probative value based on the data summarized in Table 1, it would make more sense to focus on what the data suggests about comparative propensity than on the naked recidivism rate. Table 2 contains calculations derived from the information in Table 1. Examining Table 2, one can see that prisoners released from a sentence of rape were 10.1 times more likely than the other prisoners to be re-arrested for rape, while prisoners released from burglary sentences were 2.3 times more likely than the other prisoners to be re-arrested for burglary. The figures for larceny and for drug offenses are even lower. Thus, the same-crime comparative propensity of the prisoners released from rape sentences appears to be higher than that of those released from sentences for burglary, larceny, or drug offenses. These comparative propensity statistics suggest that rape is a better candidate for an exception to a rule against other crimes evidence than those other offenses.\[128\]

---

128. I would like to note, since the point was raised during the question period at the conference with which this Symposium was linked, that the Beck study counted only offenses committed on release from custody. It did not count offenses committed in prison or after transfer to another prison. Telephone Communication from Allen J. Beck to Roger C. Park (Oct. 9, 1997); see also Beck Study, supra note 11.
Table 2. Same-Crime Comparative Propensity

<table>
<thead>
<tr>
<th></th>
<th>Ratio-</th>
<th>Ratio-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Other Offenders</td>
<td>General Population</td>
</tr>
<tr>
<td>Rape arrests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previously held for</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rape</td>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>7.70%</td>
<td>0.76%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.1</td>
<td>163</td>
</tr>
<tr>
<td>Burglary arrests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previously held for</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burglary</td>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>31.90%</td>
<td>13.70%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.3</td>
<td>56</td>
</tr>
<tr>
<td>Larceny arrests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previously held for</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larceny</td>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>33.50%</td>
<td>19.65%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.7</td>
<td>19</td>
</tr>
<tr>
<td>Drug offense arrests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previously held for</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug</td>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>24.80%</td>
<td>17.06%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>24</td>
</tr>
</tbody>
</table>

An intuition about comparative propensity may have been influential in some of the early cases that admitted other offenses of child molestation on the assumption that child sex abuse was a rare aberration. The common-sense basis for those rulings is appealing if one accepts the premise of rarity. For example, suppose a pillar of the community is accused of the crime. Investigators discover that when the defendant was employed in another city the defendant was independently accused of the same crime. On the assumption that child molestation is a rare aberration, the other crime evidence would be highly probative. Now suppose that the crime was drunk driving and an investigation turned up evidence that on another occasion the de-

129. The ratio was obtained by dividing the three-year arrest rate reported in the Beck study by three, then comparing that number with the arrest rate in the general population in 1986, as reported in FEDERAL BUREAU OF INVESTIGATION, CRIME IN THE UNITED STATES 165 (1986).

130. Cf. Lannan v. State, 600 N.E.2d 1334, 1337 (Ind. 1992) (The depraved sexual instinct exception "has its origins in an era less jaded than today. When the accusations were brought in [State v. Robbins, 46 N.E.2d 691 (1943)] in the late 1930s, the idea that an adult male who occupied a position of responsibility in the community would force himself sexually upon a child bordered on the preposterous. Sadly, it is our belief that fifty years later we live in a world where accusations of child molestation no longer appear improbable as a rule.").
fendant had driven while intoxicated. This evidence would not do as much to differentiate the defendant from the rest of the population. Moreover, it would be wiser to spend money upgrading the breathalyzer than it would be to spend it interviewing the defendant’s acquaintances in search of evidence that the defendant drove drunk on other occasions—even if the recidivism rate were higher for drunk driving than for child molestation.

I should note the limits of my own point. My assertion is that the Beck study provides no reason to question the wisdom of Rule 413. To the contrary, it provides some support for the rule. But a full examination of the question would need to go beyond what I have done here. The Beck study is only one study, though it is unusual in the size of its sample and in its same-crime comparisons. The inquiry would be furthered by a broader and deeper study of recidivism and of the incidence of crime in the general population, identifying crimes whose perpetrators have a high continuing rate of offense in comparison with the offense level in the general population.

(ii) Everybody Does It? - Rape

The concept of comparative propensity is relevant to the feminist criticisms of Rule 413 presented independently by Professors Orenstein and Baker. In part, they each rely on the argument that rape is a common crime. If many men are rapists, then that reduces the probative value of the evidence. A man’s history of rape would do less to distinguish the man from others than would be the case were rape a rare crime.

I recognize the importance of the question whether rape is common or rare. But despite the arguments advanced by Professors Orenstein and Baker, I still think that other crimes evidence should be admissible in consent-defense acquaintance rape cases. I have argued that elsewhere, and will only briefly note the main grounds of

133. Professor Baker argues against Rule 413 both on grounds that rape has a low recidivism rate and that it is very commonly committed by members of the general male population. Compare id. at 578-80 (discussing Beck study and observing that though rape may be under reported, so are other crimes) with id. at 576-77 (describing studies finding that many men rape). While it is logically possible that the incidence of rape is high and the recidivism rate for rape is low, that combination seems unlikely. If it is so, we should perhaps have renewed faith in the reformatory power of prison.
my argument here: (1) There is a special need in consent-defense cases to allow the prosecution to present evidence of other crimes (and, in cases in which it does not, to allow the defense to argue for negative inferences based on the prosecution's failure to present evidence). Otherwise the case may turn into a "swearing match" that turns on evaluation of demeanor, trivial circumstantial evidence, and the skill of the lawyers. (2) The crime is a serious one that justifies taking the time to explore situational differences. (3) The other-crime evidence is likely to be independent of the incident-specific evidence because the police have not found the defendant by "rounding up the usual suspects." (4) The danger of nullification prejudice may be less in consent-defense rape cases than in other cases, though admittedly this danger may be diminishing as public attitudes change.

Do many men rape? One way of estimating the breadth of the perpetrator base is to look at arrest rates. Here one gets relatively low numbers: a 1991 rate of 16 arrests per 100,000 population, compared to 194 for aggravated assault, 73 for robbery, 173 for burglary, and 640 for larceny-theft. Looking at offenses known to police also yields low numbers. There, one gets a rate of 42.3 per 100,000 inhabitants, compared to 273 for robbery, 433 for aggravated assault, 1252 for burglary, and 3229 for larceny-theft. Another possible source is self-report by perpetrators. Some of the surveys using this method yield high numbers. For example, a study by Mary Koss reports that 4.4% of the college students surveyed admitted to committing rape.

135. Of course, evidence that will cause error is never needed. If the danger of cognitive error were the only reason for excluding other crimes evidence, there would be nothing to be said for the argument that it ought to be admitted because it is needed. But fear of cognitive error is not the only or the strongest ground for exclusion. The dangers of time consumption, expense, and surprise are also reasons for exclusion. The danger that adversaries will neglect incident-specific evidence to pursue colorful stories of good and evil is also a reason for exclusion. These latter dangers are mitigated when there is a strong need and alternatives are missing.

136. See BJS SOURCEBOOK, supra note 11, at 423 tbl.4.2.
137. See id. at 357.
138. Mary P. Koss, Hidden Rape: Sexual Aggression and Victimization in a National Sample of Students in Higher Education, in 2 RAPE AND SEXUAL ASSAULT 3, 11 (Ann Wollbert Bengress ed., 1988), cited in Baker, supra note 125, at 576 ("In a 1988 nationwide survey of more than 6100 college students, one in twelve [sic; should be 1 in 23] college men admitted to committing rape."). See also Neil M. Malamuth, Rape Proclivity Among Males, JOURNAL OF SOCIAL ISSUES 138, 140 (1981)(35% of males surveyed in study, mostly college students, indicated at least minimal likelihood that they would rape if assured that they would not be caught).

The Koss study's 4.4% statistic for self-report of rape would have been more helpful for a judgment about comparative incidence of crimes had it been compared with self-reports of other criminal activity among college males. Koss did not ask her subjects about non-sexual crimes. The studies cited in GABOR, supra note 10, at 51-57, indicate a
Studies using survey instruments that ask about sexual coercion without using the word "rape" report even higher numbers. One can also look at victimization surveys for information about the frequency of rape. The results of one of the most thorough studies became available in 1995. The study is the redesigned version of the Bureau of Justice Statistics' National Crime Victimization Survey ("NCVS"). For this study, 100,000 individuals are surveyed annually in a continuous, nationally representative sample. The interviewers, 95% of whom are women, are census bureau employees. The interviewers make face-to-face initial contact and then follow up by telephone (unless the household has no phone, in which case they make repeated visits). In response to methodological criticisms of earlier NCVS surveys, the survey was redesigned in 1992 (with input from the National Academy of Science and the American Society of Statisticians) in order to produce more accurate reporting of sexual assault and of crimes committed by intimates or family members. The interviewers now specifically ask respondents about rape and sexual assault, encourage them to talk about sexual assault even if it is difficult, and cue respondents about victimizations by acquaintances and intimates. Of the women interviewed, 4.6 of 1,000

higher lifetime prevalence rate for other crimes. In a 1947 study of 1,700 adults from a sample selected to exclude persons with criminal records, 89% of male respondents admitted to larceny, 26% to auto theft, 17% to burglary, and 11% to robbery. See id. at 54-55. It is difficult to compare this study to Koss', however, because of differences in age (Koss' men had a mean age of 21), class (Koss surveyed college students), time period (the incidence of crime may have been lower in 1947), and social cognitions and expectations of subjects (Koss' college students may have guessed her experimental hypothesis or may have been influenced by discussion of rape in the college environment). A survey of high school seniors that was contemporaneous with Koss' study reported a one-year incidence of 30% for shoplifting, 2.8% for using a "knife or gun or some other thing (like a club) to get something from a person," and 12% for hurting someone "badly enough to need bandages or a doctor." BJS SOURCEBOOK, supra note 11, at 308-10 tbl.3.70. Again, there are obvious differences between the samples that make comparison difficult.


141. For descriptions of the methodology, see id. at 6; see also CHARLES KINDERMANN ET AL., BUREAU OF JUSTICE STATISTICS, NATIONAL CRIME VICTIMIZATION SURVEY, EFFECTS OF THE REDESIGN ON VICTIMIZATION ESTIMATES (1997); BUREAU OF JUSTICE STATISTICS, CRIMINAL VICTIMIZATION IN THE UNITED STATES, 1994 139-46 (1997).
women reported that they had been victims of rape or sexual assault (including threats) in the previous year. The rate for completed rape was 1.6 per 1000 and for attempted rape 1.3 out of 1000.\textsuperscript{142} There have been reports of much higher levels of victimization. For example, the \textit{Ms.} magazine study by Mary Koss estimated an annual victimization rate of almost 17\%\textsuperscript{143}

For the purposes of making decisions about character evidence, I think it makes sense to accept the much lower numbers indicated by the redesigned NCVS survey. It is true that there are many reasons why a woman who had been raped might not report it to an NCVS interviewer, and it seems very likely that response bias would run in the direction of under-reporting rather than over-reporting. On the other hand, it is doubtful that many women who tell an NCVS interviewer that they were not raped would tell police the opposite, or that the percentage of women ready to appear in court to testify that they were raped is higher than the percentage ready to tell NCVS interviewers that they were raped. Even if the population of rapists is large by some definitions, it seems likely that the population of men who will face multiple accusers in court will be quite small, a good deal smaller than the number of rapes reported on the NCVS survey, and a different order of magnitude from that indicated in the Koss study. At least in the case of acquaintance rape (where the accused is identified by the victim, not “rounded up” by the police from among a suspect pool of prior offenders), the men who face multiple accusers in court are quite likely from a small and particularly egregious segment of the overall male population, and hence the evidence of multiple accusations retains its probative force.

There is a second, independent reason why survey data about the frequency of rape in the general population is not conclusive. Suppose that the incidence is as high as the 4.4\% perpetrator rate for college males indicated by the 1988 Koss survey. That fact alone, while relevant and suggestive, does not tell us the frequency or level of offending by the men who have committed the crime compared to those who have not yet done so. Compare the crime of drunk driving, a crime with a large perpetrator base, as shown by both self-

\textsuperscript{142} See \textsc{Bachman} \& \textsc{Saltzman}, \textit{supra} note 140, at 5. This number is in the same ballpark as the number derived from the “offenses known to police” data. The “known to police” figure for annual rapes is 106,590 (1991) versus 313,600 for victim’s reports of attempted or completed rape (1992-93). \textit{Compare} BJS \textsc{Sourcebook}, \textit{supra} note 11, at 357 tbl.3.1222, \textit{with} \textsc{Bachman} \& \textsc{Saltzman}, \textit{supra} note 140 (1992-93 NCVS study).

\textsuperscript{143} See Koss, \textit{supra} note 138, at 13-14 (reporting results from \textit{Ms.} Magazine Project on Campus Sexual Assault; survey indicated 6-month victimization rate of 83/1,000 for rape and attempted rape). For criticism of the Koss studies, see Neil Gilbert, \textit{Miscounting Social Ills: Sexual Assault and Advocacy Research}, \textsc{Welfare Justice} 84 (1995); Neil Gilbert, \textit{Realities and Mythologies of Rape}, \textsc{Society}, May-June 1992, at 4, 9, 10.
There are major differences in the comparative propensity of those previously arrested for drunk driving and the comparative propensity of other drivers in the population. Even when a crime is common in the general population, there can be strong individual differences in propensity to offend and in the level and frequency of offending. Abstainers can be true abstainers, not merely perpetrators awaiting an opportunity.

In the example given earlier, I suggested that if all we could know was whether a driver died in a fatal accident and whether the driver had a prior history of driving under the influence, the prior history would be very helpful in deciding whether it was likely that the driver had an elevated blood alcohol level at the time of death. Of course, I do not advocate character evidence in drunk driving cases; there are better alternatives, such as tests for alcohol in the breath and the blood. But in a consent defense rape case, where better alternatives are often lacking, the prior history evidence could be very helpful.

(iii) Dangers of Using Recidivism Data

Concerns about the accuracy of the underlying data

It is often difficult to compare recidivism studies. The use of different follow-up periods and different measures of recidivism is one source of difficulty. Another one, particularly pertinent to the problem of comparing rates for different crimes, is the fact that a study of a broad category of offenders may overlook different rates of recidivism for subcategories. For example, when a study measures recidivism of sex offenders without breaking them down into subcategories, then a reportedly low recidivism rate for the group as a whole may be misleading. That could be the case if many of the per-

144. See John Mullahy & Jody L. Sindelar, Do Drinkers Know When to Say When? An Empirical Analysis of Drunk Driving, 32 ECONOMIC INQUIRY 383 (1994) (reporting that 1988 National Institute on Alcohol Abuse and Alcoholism survey found that 24% of male drinkers and 12% of female drinkers reported having driven after drinking too much within previous year; and that in 1977 and 1982 Gallup polls reported that 18% of all drivers admitted to having driven after having too much to drink at some point in life).

145. See Drivers with Repeat Convictions or Arrests for Driving While Impaired, United States, 43 MORBIDITY & MORTALITY WKLY. REP. 759 (1994) (In 1992, approximately one percent of all licensed drivers were arrested for driving while impaired.); Bureau of Justice Statistics, Sourcebook-1992 at 423 (arrest rate in 1991 for driving under the influence of 678.5 per 100,000).

146. See infra text accompanying notes 21-23.

147. See id.

sons studied had been convicted of intrafamilial sex abuse and had a low recidivism rate because of dynamics specific to that crime, while the extrafamilial child abusers in the same study had a hidden high recidivism rate. There are also dangers of selection bias, which challenge the generalizability of individual studies and make comparison of studies difficult: some studies examine hard-core offenders who have multiple offenses; others examine offenders who have been admitted to treatment programs that require them to accept responsibility and to desire change. As one social scientist noted, by selectively contemplating the studies of recidivism of sex offenders, one can argue for almost anything.\textsuperscript{149}

Data based on re-arrest rates are vulnerable both to false positives (innocent persons known to have a record for the crime may be arrested on suspicion, then released) and to false negatives (offenders may repeat many times without being re-arrested).\textsuperscript{150} The problem of undetected offenses is a daunting one. One study that surveyed offenders in a prison and treatment center received reports from them that they had committed, on average, 5.2 undetected rapes (after excluding those who claimed 50 or more undetected offenses).\textsuperscript{151} Studies of this nature have their own potential flaws: the most obvious of

\begin{itemize}
\item \textsuperscript{149} See id. at 27 (quoting Vernon L. Quinsey, Sexual Aggression: Studies of Offenders Against Women, in 1 LAW AND MENTAL HEALTH: INTERNATIONAL PERSPECTIVES 84, 101 (D. Weisstub ed., 1984).
\item \textsuperscript{150} Many experts consider arrest records to be a better measure than records of prosecution or conviction. Though use of arrest records could inflate recidivism rates—for example, known child molesters might be arrested on suspicion even though they have done nothing—most experts consider the arrest measure to be superior to other measures based on law enforcement records. Relying upon conviction records would exclude repeat offenders who are put into diversion programs, granted immunity in return for testimony, or not prosecuted simply because the prosecutor has more important cases. Also, for studies that aim at ascertaining whether the perpetrator committed the same offense again, use of conviction records would lead to misleading results when the perpetrator is allowed to plead to a lesser offense (though of course a tendency of police to overcharge may lead to misleading results when arrest records are used). For discussion, see MALTZ, supra note 119, at 54-58.
\item \textsuperscript{151} See A. Nicholas Groth et al., Undetected Recidivism among Rapists and Child Molesters, 28 CRIME & DELINQ. 450, 453-54 (1982). The researchers surveyed 83 rapists and 54 child molesters. The child molesters also reported undetected criminal activity. Because of the fact that a child molester will often commit repeated acts with the same victim, the researchers did not attempt to count the actual number of criminal acts, but only the number of separate victims. On average, the child molesters reported undetected crimes involving 4.7 victims. See id.
\end{itemize}
which is relying upon the accuracy of offenders in responding to survey questions, and concerns about whether the samples surveyed are representative of the population of offenders. However, they do strongly suggest that one should not be too eager to rely upon re-arrest or re-conviction data, and that the true recidivism rate may be practically unknowable.

**Difficulty of drawing conclusions even from accurate data**

Recidivism data is not gathered for the purpose of informing evidence scholars about the probative value of character evidence. Data that is valuable for other purposes—such as aiding decisions about preventive detention or evaluating rehabilitation programs—may be less valuable for drawing conclusions about other-crimes evidence, for the following reasons:

(a) Recidivism is not “precidivism”

Recidivism studies measure the rate at which convicted offenders recommit an offense. But reforms such as Rules 413 and 414 of

---

152. For example, in the Groth survey reported above, one subject reported 250 undetected sexual assaults. *See id.* Under the rules of the study, this would mean that the subject either victimized 250 separate children or committed 250 separate adult rapes that were not detected. Seven percent of the subjects reported 50 or more undetected sexual assaults. These high-incidence subjects were excluded by the researchers from their calculations, but it is possible that other respondents made exaggerated reports on the anonymous survey. *See id.* at 454-55. On the other hand, respondents may well have understated their sexual activity. Groth et al. believed that this was probably the case, stating that “[i]t is more characteristic for sexual offenders to minimize their wrongdoing than it is for them to exaggerate their criminal activity.” *Id.* at 456.

153. In summarizing their data in approximate terms, Groth et al. state that “[t]he rapists, who had an average of three rape convictions on record, and the child molesters, who averaged two convictions on record, admitted to an average of five similar offenses for which they were never apprehended.” *Id.* at 456. One might wonder whether a group that had “three rape convictions on record” might not be a hard-core group with atypical experiences.

154. In addition to the Groth et al. study cited above, see David Finkelhor, *Abusers: Special Topics, in A SOURCEBOOK ON CHILD SEXUAL ABUSE* 132 (David Finkelhor ed., 1986). Finkelhor noted that ten studies of child molestation “probably gravely understate the amount of subsequent offending committed by the men who were studied. The investigators routinely used as their criteria of recidivism subsequent offenses that came to the attention of the authorities.” *Id.* See also Judith V. Becker & John A Hunter, Jr., *Evaluation of Treatment Outcome for Adult Perpetrators of Child Sexual Abuse*, 19 CRIM. JUST. & BEHAV. 74, 82 (1992)(stating that “undetected crime is quite extensive among sex offenders and... official data may reveal only a small percentage of the total sexual offenses committed”); Thomas J. Reed, supra note 120, at 150 (suggesting that the “low visibility of sex offenders in general and rapists in particular obscures a high recidivism rate for rapists”).
ten make evidence of prior offenses that did not result in conviction admissible. Evidence of prior crimes and bad acts can be offered in cases in which the defendant, though possibly guilty of other crimes, had never been convicted or charged with any crime. In such cases, data about re-commission of crimes by persons who have been convicted is relevant only to the extent that it sheds light on the continuity of behavior of persons who have never been convicted. On the issue of guilt the trier must decide a point of historical fact, not make a prediction about what the defendant might do in the future. A well-designed recidivism study that would be helpful in evaluating treatment programs or in deciding issues of sentencing policy might for this reason not be very helpful in weighing the probative value of prior crimes in determining an issue of historical fact.

Behavior patterns can change as a result of incarceration, treatment, or other conviction-related consequences. To take an extreme example, the reported recidivism rate of sex criminals who have been castrated is quite low (though it does not drop to zero);\textsuperscript{155} but obviously one could not use that recidivism data to support inferences about how often the offenses were repeated before the offenders were apprehended. Similarly, one study reported that defendants convicted of intrafamilial incest had only a 4\% recidivism rate for any sex offense in the 12 years after conviction. However, that result can be explained with the hypothesis that the offenders took advantage of opportunity within the family, and that once their crimes were discovered, family pressure (or removal from the family) prevented further crimes from being committed.\textsuperscript{156} Age alone is a very important factor, especially in light of the length of modern American sentences for serious crimes. It is not surprising to find the recidivism rate for convicted murderers to be low, if only because their productivity as murderers is likely to be impaired by age by the time they are released.\textsuperscript{157} (Indeed, Gottfredson and Hirschi advance the age effect as a reason why selective detention must fail—by the time the subject has accumulated enough criminal history, age has changed him.\textsuperscript{158}) In considering character evidence, of course, one could still

\textsuperscript{155} See Furby et al., \textit{supra} note 148, at 17.

\textsuperscript{156} See \textit{id. at} 5.

\textsuperscript{157} See BECK STUDY, \textit{supra} note 11, at 5 tbl.7 (recidivism inversely related to age of prisoner at the time of release: the older the prisoner, the lower the rate of recidivism).

\textsuperscript{158} See GOTTFREDSON & HIRSCHI, \textit{supra} note 10, at 262-63.
expect that the evidence might be useful because the criminal propensity of a middle-aged accused with a long criminal history will be much higher than the criminal propensity of a middle-aged person chosen at random. But in comparing recidivism data for what it says about propensity across crimes, age is a potential confounding variable because some crimes result in much longer sentences.

As the examples suggest, when making inferences about the "precidivism" rate of individual crimes from recidivism data, one needs to consider the possibility that different crimes have different recidivism/precidivism ratios. For example, imprisonment could be criminogenic for burglars, while having a deterrent effect on sex offenders. This could be the case if burglars leave prison at an earlier age than sex offenders because of shorter sentences (thus reducing maturation effect), if they learn new techniques in prison, or if they return to burglary after release because their criminal records foreclose other employment.

(b) Knowledge of recidivism influences actors in the criminal justice system

Another obstacle to straightforward reliance on recidivism data is that prior crime evidence may play a role in causing suspicion to focus on the defendant in the first place. The police may focus on prior offenders when looking for a burglar, robber, or assailant. They may, subtly or blatantly, encourage a victim to identify someone whom they suspect on the basis of knowledge about the suspect's prior offenses. This "round up the usual suspects" danger complicates the use of recidivism information. If the police use formal or informal knowledge of recidivism as a predictor of repeated crime, then they are likely to arrest past offenders more often for crimes having high recidivism rates than for crimes having lower rates. In any event, police might select prior offenders with records of serious crimes more frequently than prior offenders with records for minor crimes.

b. Seriousness

Another criteria that could be considered under a crime-by-crime approach is the seriousness of the offense charged. A more serious offense justifies a more costly trial. The importance of the decision might justify going into the details of prior crimes; examining carefully the similarities and differences in situations, motives, expectations, and goals; hearing the testimony of (and the ancillary testimony impeaching and supporting) witnesses who maintain that prior events did or did not happen; and perhaps even having experts con-
duct repeated interviews and offer clinical judgment. Thus, serious crime is a natural candidate for exemption from the character ban.

The problem, however, is that just as seriousness justifies cost, so too does it increase the chance of nullification. Serious crimes are likely to be horrifying. If the prior crimes are also horrifying (as would be the case in a crime-by-crime approach that allowed crimes with a similar statutory definition to be admitted) then the danger of nullification prejudice increases. While seriousness is an important factor, it cuts both ways to such an extent that it does not help much in deciding which crimes to choose.

c. Need

Reformers should also consider the need for the evidence. Admittedly, taking need into account would make no sense were prejudice the sole basis for exclusion. There is never a need for evidence that will lead to bad decisions. But if cost is a large factor in the rationale for exclusion, then need becomes very important. When there are better alternatives to the exploration of character (as in the drunk driving situation, where drunkenness can easily be tested) then the need for character evidence, and hence the justification for paying its cost, disappears.

The problem with considering need as the basis for a crime-by-crime approach is that need depends more on the particular circumstances of a case than on the type of crime. That is not to say that there are not types of crimes where the need may be greater. Consent-defense rape cases, which often boil down to a credibility contest between two witnesses, may be one example. But in general it is hard to say whether the need for character evidence is categorically greater in murder cases than in, say, arson cases. Need cannot be assessed without a case-specific knowledge of the alternatives.

(3) Disadvantages of a Crime-by-Crime Approach

The crime-by-crime approach might seem a natural reaction to the precedent set by the sex crimes legislation. The approach could make the law seem a bit more consistent, for example by treating all violent felonies the same way, or simply by adding homicide to the list of crimes for which character proof is allowed. Abolition of the character evidence ban for homicide would allow a thorough airing of the facts in this most serious crime. It would also, however, open up character proof in the area where the danger of nullification prejudice is at its apex.

There are other problems with the approach. By adding specific-crime exceptions to the existing rule structure, the crime-by-crime approach would retain old complications while creating new ones.
The enigmas of Rule 404(b) and 405(a) would persist for crimes not covered by the new exceptions, and the new exceptions would themselves be complex, assuming they followed the model of Rules 413 and 414.159

The crime-by-crime approach also muddles the principles supporting the character evidence rule. The character evidence ban helps teach that the question is whether the defendant committed the acts charged, not whether the defendant is a good or evil person. The rule's effectiveness as a moral teacher is diminished when some crimes become admissible for character purposes. Of course, the jury can still be instructed to consider the character evidence only for its bearing on the acts charged, and not to punish for other evil done. But if the prosecution can argue that the defendant is a rapist or murderer and tell a colorful story of evil supported by specific instances, then the message may be harder to convey. Further confusion will be created by the jury instructions in multiple-offense cases in which some crimes are admissible for character reasoning and others are admissible only on Rule 404(b) grounds.

As noted earlier, I believe that obstacles to admitting other crimes evidence should be lowered in consent defense rape cases. I would, however, accomplish that purpose simply by adding "mistake as to consent" and "sexual coercion" to Rule 404(b)'s KIPPOMIA list. The change could lead to some messy jury instructions ("use the evidence only to show sexual coercion, not bad character") but the instructions would be no more messy than the ones now given in cases in which KIPPOMIA evidence is offered to show intent. One more little stone in the grotesque edifice might prevent the injustice that has occasionally occurred when courts have rejected evidence to show sexual coercion that would have been admissible if offered to show identity.160 Of course, Rule 403 would still be available to defendants complaining of the admission of evidence of crimes that were remote or highly dissimilar.

159. Federal Rules of Evidence 413 and 414 do not apply to all sex crimes. The drafters particularized in order to create better evidentiary matches. They described in detail the crimes that could be used for character reasoning, sometimes making cross-references to criminal code sections. The result was an intricate rule structure that even evidence professors cannot recite from memory.

160. KIPPOMIA is an acronym, coined by Richard Uviller, that refers to knowledge, intent, plan, preparation, opportunity, motive, identity, and accident. See Uviller. supra note 34, at 877. Cf. Fed. R. Evid. 404(b).

B. Structured Balancing

(1) Examples of Structured Balancing

In an attempt to address issues of probative value with greater discrimination, some commentators have advocated an approach that I will call "structured balancing." By "structured balancing," I mean to refer to proposals to allow evidence of other crimes and wrongs to be used for character reasoning, subject to certain requirements such as frequency, proximity, and similarity. In contrast to the KIPPOMIA approach, the structured balancing approach does not attempt to identify a titular purpose other than character reasoning for which evidence of other crimes can be used.

Professor Uviller’s proposal is an example of structured balancing. The part of his proposal that relates to substantive use of character evidence (stated in its essence, with some elaborations omitted) would provide that "evidence of character shall be admissible" in the form of (1) "opinion, but only as to the trait of rectitude or other traits narrowly defined and modified by circumstances closely resembling the situation in issue; [and] (2) specific instances of prior conduct, but only by repeated, habitual, or unusual incidents of similar behavior in similar circumstances."163

In her article on psychology and character evidence, Susan Davies also offered suggestions for doctrinal reform. She favored guided discretion, with attention to specificity of conduct, similarity, frequency, duration, and "whether the conduct is inherently likely to be repeated."164

The Judicial Conference has reluctantly provided a similar model. When faced with the likelihood that the Rules 413, 414, and 415 would create a blanket exception to the character ban in sex crime cases, the Conference came up with a plan for structured balancing aimed at those types of cases, as a second choice to keeping the existing system and not allowing character evidence at all. Its

162. See Uviller, supra note 34, at 889.
163. Id.
164. Davies, supra note 29, at 535-36. She states that:
The psychological studies favor discretionary standards that require the judge to have regard, as appropriate, to factors such as the specificity of the description of the subject’s prior conduct, the similarity between that conduct and the surrounding circumstances in which it occurred and the conduct in question, the number of provable prior instances of similar conduct by the defendant, the period over which the prior conduct occurred, and whether the conduct is inherently likely to be repeated.

Id.
The proposal would have mandated that in sex crime cases, the judge consider factors such as proximity in time, frequency, similarity of situation, and intervening events.\textsuperscript{165}

The structured balancing proposals described above aim at identifying factors that enhance the probative value of the other crimes evidence, and echo some of the factors derived from personality theory. These proposals would be enhanced by adding language that takes into account the adversarial context, for example mandates to consider need for the evidence and the alternative proof available. They could also be supplemented with notice provisions and a Rule 403-like authority to consider cost and confusion.

(2) \textit{Evaluation of Structured Balancing}

Structured balancing is an attractive alternative. It points the trial judge toward factors identified in psychology scholarship as relevant to regularity of behavior. It seems more likely to encourage meaningful on-the-record balancing than does current Rule 404(b).

On the other hand, the three proposals I have described would probably not reduce the complexity of the system of rules. Structured balancing proposals that admit "character evidence" upon a showing of frequency and similarity would need to be supplemented by a rule that bad acts evidence is admissible when it is for a purpose other than showing character. Otherwise, single-act evidence that clearly ought to be received would have to be excluded. Though Professor Uviller states that his system would make the KIPPOMIA list obsolete, it is hard to see how this could be so. Consider the following examples:

\begin{quote}
165. Here is the Judicial Conference's fall-back rule:
[Rule 404 (a)](4) Character in sexual misconduct cases. Evidence of another act of sexual assault or child molestation, or evidence to rebut such proof or an inference therefrom, if that evidence is otherwise admissible under these rules, in a criminal case in which the accused is charged with sexual assault or child molestation, or in a civil case in which a claim is predicated on a party's alleged commission of sexual assault or child molestation. (A) In weighing the probative value of such evidence, the court may, as part of its rule 403 determination, consider: (i) proximity in time to the charged or predicate misconduct; (ii) similarity to the charged or predicate misconduct; (iii) frequency of the other acts; (iv) surrounding circumstances; (v) relevant intervening events; and (vi) other relevant similarities or differences

(b) Other crimes, wrongs, or acts. Evidence of other crimes, wrongs, or acts is not admissible to prove the character of a person in order to show action in conformity therewith except as provided in subdivision (a).

\end{quote}
Table 3. Single Act Evidence Admissible Under Existing Law

<table>
<thead>
<tr>
<th>Prior crime</th>
<th>Charged crime</th>
<th>KIPPOMIA category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stealing a key</td>
<td>Burglarizing a safe with the stolen key.</td>
<td>Plan, preparation, identity</td>
</tr>
<tr>
<td>Stealing a license plate</td>
<td>Robbing a bank in a car bearing the stolen plate</td>
<td>Plan, preparation, identity</td>
</tr>
<tr>
<td>Any crime, observed by Y</td>
<td>Killing Y to eliminate a witness</td>
<td>Motive, identity</td>
</tr>
<tr>
<td>Sneaking into a guarded off-limits area on day one</td>
<td>Theft from the area on day ten</td>
<td>Plan, preparation, opportunity, identity</td>
</tr>
<tr>
<td>Smoking marijuana</td>
<td>Growing marijuana, claiming not to know what it was</td>
<td>Knowledge</td>
</tr>
</tbody>
</table>

I have no doubt that after structured balancing reform, the prior crimes listed above would still be admitted, even if they did not fit the structured balancing test. In doing so, the judge admitting the evidence (or sustaining its admission) would be likely to say that the structured balancing test does not have to be satisfied, because the evidence is not offered to show character, but to show something else: knowledge, intent, plan, preparation, opportunity, motive, identity, or absence of mistake or accident. Old rules die hard.

The other problem with structured balancing is simply one of cost and time-consumption. The greater the similarity and frequency, the greater the chance of admission. But knowledge about similarity and frequency does not just drop into the trier's lap. It must be explored through witnesses who may disagree and whose credibility may raise further side issues. Of course, the greater the alleged frequency, the greater the number of side issues.

C. Free Balancing

The rules of character evidence are so frustrating that it is tempting to give up on them completely, resorting to what I will call "free balancing"—that is, the Rule 403 approach, perhaps with the scales reversed. Rule 403 does not try to describe situations, but merely lists purposes, such as avoiding cost, confusion, and prejudice. We see examples of this approach in proposals by Professors Kuhns\textsuperscript{167} and Leonard.\textsuperscript{168}

\textsuperscript{166} FED. R. EVID. 403.
\textsuperscript{167} Kuhns tentatively proposed a rule that "might provide":

HeinOnline -- 49 Hastings L.J. 777 1997-1998
In assessing such a proposal, one must first ask whether meaningful appellate review would be retained. If so, then one is likely to see KIPPOMIA factors creeping back in as the appellate courts seek to explain their rulings. After all, it is hard to resist saying that evidence that the defendant stole the key to open the safe is not offered to show general bad character, but merely to show opportunity and plan. That is a pretty sensible way of pointing out that in the case at bar, the character dangers are not present. Again, the old rules do not go away that easily.

If, however, meaningful appellate review is not intended—if trial judges are to have unfettered discretion—then of course free balancing would cause a sea of change, at least with newer judges not so strongly imbued with the past system of rules and exceptions. But then it would give judges a higher degree of discretion than we have been willing to tolerate in this area of the law. The multitude of cases on other-crimes evidence are both a healthy and unhealthy sign. Their number shows that the ban is hard to administer, but also suggests a genuine appellate attempt to keep trials fair. One can easily imagine great unfairness at the hands of a biased judge. One party might, for example, be allowed to introduce bad acts while the other was not, leaving the jury to mistakenly think that only one party had stains on character. The power to exclude and admit without oversight is the power to tailor a result, help friends, punish enemies, and generally undermine the right to trial by jury.

A balancing approach, whether relatively structured or relatively free, would avoid the inconsistencies and rigidities of the crime-by-crime approach. It might lead to more meaningful on-the-record consideration of relevant factors than is the case under current Rule 404(b) practice. The problem is one of controlling discretion. Perhaps it would help to enhance requirements for case-specific justification of decisions as an offset to relaxation of requirements that evidence fit within a preordained category of proof. Judges could be required to justify the admission of other-crimes evidence (at least in close cases) by an extensive discussion of its probative value and its linkage to case-specific evidence. Were such a regime really enforced by appellate courts, then perhaps the bother of writing such findings

Except as otherwise provided by Act of Congress or by these rules, evidence of other crimes, wrongs, or culpable acts is not admissible for any purpose unless the court determines that the probative value of such evidence outweighs the danger of prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.


alone would be enough to deter reckless reception of character evidence.

Conclusion

Readers who have made it this far have noticed the ebb and flow of an author arguing with himself, the back and forth of someone who is skeptical of other solutions but not sure of his own. Commentators who favor the character evidence ban have exaggerated its value in preventing inferential error and have overstated the degree to which scholarship on personality supports the ban. Nonetheless, character evidence is undeniably dangerous. Perhaps its dangers could be reduced by reform of other aspects of the criminal justice system, but those reforms do not seem to be on the horizon, and it is doubtful in any event that they would be packaged with character evidence reform.

One thing is certain: complexity will be with us for a long time. It is inherent in the project of inference regulation by rule. So long as we try to control both prejudice and discretion, other-crimes cases will be as common as cars in a city.  

169. McCormick compared the other crimes cases to sands of the seas, see McCORMICK ON EVIDENCE 190, at 558 n.8 (Edward W. Cleary ed., 3d ed. 1984), but I prefer cars in a city, on grounds that they are humanly constructed and give off noise and pollution.