Myself Alone: Individualizing Justice Through Psychological Character Evidence

Andrew E. Taslitz

Follow this and additional works at: http://digitalcommons.law.umaryland.edu/mlr
Part of the Evidence Commons

Recommended Citation
Available at: http://digitalcommons.law.umaryland.edu/mlr/vol52/iss1/4
Articles

MYSELF ALONE: INDIVIDUALIZING JUSTICE THROUGH PSYCHOLOGICAL CHARACTER EVIDENCE

ANDREW E. TASLITZ*

Table of Contents

I. Psychological Character Evidence and the Attribution of Responsibility .................. 6


1. Criminal Acts ................................................. 9

2. Guilty Minds (the Susceptibility Defenses) .......... 10

   a. Negating Mental State ................................. 10

   b. Mitigation ................................................ 11

   c. Affirmative Defenses ................................. 12

B. Criminal Justice and the Psychology of Moral Attribution ...... 14

   1. Rule Logics ................................................. 14

   2. Case Logics ............................................... 16

   3. Typification and How to Overcome It ............... 18

C. Morality and the Principles of Criminal Evidence ........ 20

D. Individualized Justice and the Spectrum of Typification ...... 24

   1. The Upper End of the Spectrum: The Group ........ 25

* Associate Professor, Howard University School of Law; former Assistant District Attorney, Philadelphia, Pa.; B.A., Queens College, 1978; J.D., University of Pennsylvania Law School, 1981.

The author thanks his wife, Patricia V. Sun, Esq., and Professors Laird Kirkpatrick and Jose Alvarez for their numerous helpful comments on earlier drafts of this Article. The author also thanks Dr. Ellen McDaniel for her comments on many of the scientific issues addressed herein. Appreciation additionally goes to the author's research assistants, Alfred English, Julie Abbate, Jordan Gross, and Cheryl Moat for their help in completing this Article, and the Howard University School of Law for its financial support of this project.
a. Probabilistic Prediction .......................................................... 26
b. Correcting Misconceptions .................................................... 27

2. The Lower End of the Spectrum: The Individual .................... 29

II. TOOLS OF THE TRADE .......................................................... 30
A. The Concept of Personality .................................................... 31
B. Objective Personality Tests .................................................... 34
   1. Creation and Measurement ................................................. 34
   2. A Brief Comment on the MMPI ............................................ 37
C. Projective Tests ................................................................. 39
D. Observations and Interviews ................................................ 39
E. Diagnoses and DSM-III-R ..................................................... 41
F. Profiles, Syndromes, and Typologies ..................................... 42
G. Case Studies and Life Stories ............................................... 43

III. EVIDENTIARY ANALYSIS ...................................................... 44
A. Federal Rule of Evidence 404 ................................................ 44
B. The Frye Rule ........................................................................ 48
   1. Is Expert Character Testimony "Scientific"? ......................... 51
   2. Is Expert Character Testimony "Novel"? .............................. 56
   3. Are "Overt" and "Covert" Group Comparisons the Same? .......... 57
   4. Is the "Cathartic Function" Served by Frye? ....................... 60
C. Pragmatic Relevancy and Psychological Character Testimony 63
   1. Reliability ........................................................................... 64
      a. The Predictive Power of Traits ....................................... 65
         i. The Principle of Aggregation ....................................... 66
         ii. Average Behavior ...................................................... 67
         iii. Improving Predictive Power ...................................... 69
      b. Experimental Data versus Clinical Judgment .................... 72
         i. Experimental Data ...................................................... 72
         ii. Clinical Judgment ...................................................... 76
      c. Witness Qualifications .................................................... 81
      d. Negative Inferences ....................................................... 83
   2. Necessity ............................................................................. 86
      a. The Political Role of the Jury .......................................... 86
      b. Suppositional Science, Plausible Alternatives, and Storytelling ......................................................... 91
         i. Suppositional Science as a Way of Suggesting Plausible Alternatives to Normal ........................................ 91
         ii. Storytelling Theory ...................................................... 94
      c. Art versus Science .......................................................... 98
         i. Why Art is Helpful ...................................................... 98
         ii. "Good" Art versus "Bad" ............................................... 100
      d. Availability of Less Objectionable Alternatives ............... 102
The first thing the intellect does with an object is to class it along with something else. But any object that is infinitely important to us and awakens our devotion feels to us also as if it must be *sui generis* and unique. Probably a crab would be filled with a sense of personal outrage if it could hear us class it without ado or apology as a crustacean, and thus dispose of it. "I am no such thing," it would say; "I am MYSELF, MYSELF alone."¹

— William James

The cry to be judged for "myself alone" is heard daily in the criminal courtroom.² This is because most laws governing criminal conduct require an inquiry into the mental state of the accused.³ The question is not what most reasonable persons would think but rather what the defendant in fact thought when allegedly committing the crime.⁴ Mental state determines whether a defendant is guilty of a crime and, if so, the degree of the crime. For example, homicide may be punished less severely when the killer is merely aware of but ignores a risk of death—"recklessness"—than when the killer desires the victim's death—"purposefulness."⁵ The defendant’s unique thoughts thus govern the degree of culpability.

An individualized inquiry is also required in proving acts. Unlike civil cases, where the identity of the parties is generally known, the central determination in criminal cases is often discovering who did the wrongful deed.⁶ Consequently, the jury’s focus must be on

---

2. See infra text accompanying notes 56-118.
4. See Morissette v. United States, 342 U.S. 246, 250 (1952) ("The contention that an injury can amount to a crime only when inflicted by intention is no provincial or transient notion. It is . . . universal and persistent in mature systems of law . . . ."); JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 102 (1987) ("Punishment without evidence of subjective fault (i.e., punishment for negligence) is highly controversial.").
5. See MODEL PENAL CODE §§ 2.02(2)(a),(c), 210.2, 210.3 (defining recklessness and purpose and assigning more severe punishments to purposeful killings than reckless ones). Note, however, that a homicide involving a special type of recklessness, that is, recklessness showing an extreme indifference to the value of human life, is punished just as severely as homicide involving purpose. See id. at § 210.2.
whether a particular defendant committed the crime. In making this judgment, what the defendant thought may be highly relevant in proving what she did, so again an examination of what was in the defendant’s mind—not what should have been there—is key.7

The law thus requires that each defendant be treated as unique, a “universe of one.”8 But what the law requires and what the defendant receives are often two very different things. Overburdened criminal courts behave more like bureaucracies than centers of justice.9 The pressure to “move cases” leaves little time, and fewer resources, to treat each defendant as unique. Add to these pressures the stereotypes, myths, and presuppositions that we all bring to the courtroom and the result is trial by assumption, not by a fair and thorough evaluation of each defendant as a special human being.10

The question, therefore, is whether there are ways to move judges and juries closer in practice to the “individualized justice” required by the basic tenets of criminal law. This Article suggests one way to help achieve that goal: using expert character testimony. Because such testimony focuses on what “kind of person” the defendant is, it can tell the jury the defendant’s story and of special circumstances that may require treatment different from the “average person.”11 Correspondingly, the prosecution, although more rarely than the defense, may seek to focus on the ways in which the defendant differs from the average. In this way, expert character testimony, or, to reflect more precisely the type of experts involved, “psychological character evidence,”12 offers the prospect of moving the jury toward the individualized assessments of guilt that the law requires.

tifying the defendant as the wrongdoer presents an issue, and often the sole one for determination, in every criminal trial.”).

7. See infra text accompanying notes 83-128, 136-141.
8. DONALD SCHON, THE REFLECTIVE PRACTITIONER: HOW PROFESSIONALS THINK IN ACTION 108 (1983) (coining the term “universe of one” to describe how psychotherapists view their patients).
9. See infra text accompanying notes 83-118.
10. See infra text accompanying notes 83-118.
11. See Juanita R. Brooks, Creative Defenses and Desperate Defenses, 18 LITIG. 22 (Winter 1992) (making a similar point from a practitioner’s perspective, although the author did not explicitly recognize that much of what she described in her article was “psychological character evidence”); infra text accompanying notes 83-141.
12. The term “psychological character evidence” is used as a shorthand for expert opinions regarding character, regardless of whether the expert is a psychologist, psychiatrist, or other qualified evaluator.
But how good are psychologists in making these kinds of judgments? Can the jury understand and effectively use expert character testimony without being "overawed" by it? Questions of this type, which focus on whether the probative value of the evidence is outweighed by the danger that it will mislead the jury—what some have called "pragmatic relevancy"—are the subject of this Article. Consequently, the Article examines much of the science underlying expert character evidence to develop a sense of its probative value. It also examines psychological research regarding the effect of such evidence on a jury.

In weighing the balance between the benefits and burdens of using psychological character evidence, this Article argues that attention must be paid to the moral and political bases for the rules of evidence, rather than focusing solely on how those rules affect truth-seeking. These bases are too often ignored in evidence scholarship and case law. This Article further argues that evidentiary rules need not be used primarily as screening devices to weed out inadequate evidence. Instead, evidentiary rules can also be used to increase the quantity and quality of the evidence available to the jury.

The Article concludes that the powerful role that expert character evidence can play in individualizing justice requires trial judges to adopt an activist stance that will enable them to improve the value of such evidence to a jury in a fashion that recognizes the moral and political values involved and the sharp differences between the laws of evidence in civil and criminal trials.

This Article is also about stories, including stories that tell what expert character witnesses do and how they do it. The Article recounts some stories told by such witnesses, and examines recent theories maintaining that the creation of stories that make sense is precisely what jury deliberations are designed to produce. To the extent expert character testimony helps to build better stories, it will be of particular value for both prosecutors and defendants in explaining facts to jurors in a way that will make the most sense to them.


Part I of this Article, "Psychological Character Evidence and the Attribution of Responsibility," examines the moral and psychological bases for the criminal law's requirement of individualized justice and the social forces that make that goal difficult to achieve. It is precisely these realities, this part argues, that require treating the principles of criminal and civil evidence very differently. Part I concludes with an overview of how psychological character evidence can help overcome bureaucratized trial by myth, stereotype, and supposition, focusing jurors' attention instead on the person before them.

Part II, "Tools of the Trade," examines the techniques and instruments expert character witnesses use in plying their trade and how they apply them. This review of both the science and the art of the psychologist, the typical expert character witness, offers the reader an understanding of the kind of information that can be conveyed to a jury to help it evaluate the evidence before it. This section also offers information that may be useful to practicing lawyers in preparing direct and cross examinations of expert character witnesses.

Part III, "Evidentiary Analysis," answers the pragmatic relevancy-related objections to psychological character evidence. Moral, scientific, and artistic principles are combined with a study of the political role of the jury, to craft a set of guidelines for courts in addressing many admissibility questions concerning expert psychological character evidence. The guidelines emphasize, contrary to the usual focus of American evidence law, the value of persuading and cajoling, not commanding, jurors into doing what is right.

I. PSYCHOLOGICAL CHARACTER EVIDENCE AND THEATTRIBUTION OF RESPONSIBILITY

The argument in this part proceeds in three steps. First, moral psychology teaches us that the human sense of justice requires individualized treatment in criminal cases, yet powerful social forces

15. "Pragmatic relevancy" concerns are usually addressed under the rubric of Federal Rule of Evidence 403, which permits a court to exclude evidence if its probative value is substantially outweighed by certain countervailing concerns, particularly improper jury impact. See id. at 79. There are, however, other evidentiary rules that have a weighing of probative value against jury impact as their primary concern. These rules are therefore also addressed in this Article. Furthermore, this Article argues for an expanded notion of pragmatic relevancy in which the judge's discretion under Rule 403 is guided in part by moral, political, and fairness-of-process concerns, concepts not usually addressed in standard ideas of pragmatic relevancy. There are many non-pragmatic relevancy issues raised but not addressed herein (for example, hearsay-related and constitutional questions), but these issues will have to be addressed in future articles.
push juries away from such treatment. Second, our formal rules of criminal law embody this important but, in practice, unfulfilled psychological need; to remedy this inadequacy, we must recognize that the law of criminal evidence can help to return the justice system to an individualized ethic by increasing the flow of quality information to the jury. Third, psychological character evidence is one means by which we can, in appropriate cases, effectively move the system toward greater individualization. Before reaching these points, however, it is first necessary to define psychological character evidence and illustrate it more clearly.

A. What Is “Psychological” Character Evidence?

“Character evidence,” as the term is used in this Article, is evidence offered to help the jury answer the question, “What type of person is the accused?” It is any testimony by a psychologist or other mental health professional that suggests that because a defendant is a certain kind of person, she is less likely—or more likely, in the case of prosecution testimony—to have committed the criminal deed with the requisite level of intent. To be useful, such evidence must ordinarily concern a specific pertinent trait of the accused, not a general description of how she is likely to think or act in all situations. Because psychological character evidence rests upon the expertise and training of a mental health professional, it must be offered in the form of that professional’s opinion, based upon examination and testing, concerning a particular trait of the accused’s personality.

Note that this definition, contrary to the definition of “character” offered by some commentators, does not necessarily have a moral connotation: the testimony need not concern whether the de-

16. See CHARLES T. MCCORMICK ET AL., MCCORMICK ON EVIDENCE § 186, at 549-50 (Edward E. Cleary, ed., 3d ed., 1984). The rules, of course, permit the use of character evidence as to persons other than the accused. See FED. R. EVID. 404 (a)(2)-(3), 608 (permitting certain uses of character evidence to prove a pertinent trait of a crime victim and the character for truthfulness or untruthfulness of a witness). However, this Article focuses solely on evidence as to the character of the accused.

17. See MUELLER & KIRKPATRICK, supra note 14, at 487 (noting that character evidence is most significant when it concerns specific inclinations); FED. R. EVID. 404 (a)(1),(2) (requiring defense-proffered character evidence to concern a “pertinent” trait of the accused).

fendant is in some sense a good or a bad person.19 This makes sense because a conception of character evidence based solely on morality would be inconsistent with the policy concerns that have led courts to treat such evidence cautiously.20 Courts have often been wary of character evidence, fearing that juries will convict or acquit because of who the defendant is—a good or bad person—rather than what the defendant has done.21 The law, however, also has broader concerns.

Specifically, when character evidence is used circumstantially, a jury must engage in a two-part analysis: it must determine, first, whether the evidence offered as to the defendant’s reputation, opinion, or specific acts proves that the defendant indeed has a certain personality trait and, second, whether someone with that trait is more—or less—likely than someone without it to commit the criminal deed. The value of step one turns on the trustworthiness of the proffered evidence: whether it suggests that the defendant has the identified trait. The value of step two turns on another accuracy judgment: whether having the trait affects the likelihood that the defendant committed the crime. The jury may, without proper guidance, overvalue the worth of the evidence on both points. That is why courts so mistrust character evidence.22

These dangers exist with much evidence that does not necessarily carry moral overtones about a defendant’s worth as a human being. Indeed, the drafters of the Federal Rules of Evidence rejected the traditional criticism that character evidence is primarily offered for “moral overtones,” noting that “nonmoral considerations [are] cropping up” more frequently and describing the proper focus as “the kind of person one is.”23 The drafters specifically offer the example of a psychiatrist’s opinion—what would be identified here as psychological character evidence—as a way to prove that someone is

---

20. See infra text accompanying notes 244-265.
22. See Imwinkelried, supra note 19, at 575-84.
23. See FED. R. EVID. 405 advisory committee’s note (“If character is defined as the kind of person one is, then account must be taken of varying ways of arriving at the estimate.”).
indeed a particular "kind of person."\(^{24}\)

This definition, and the practical value of this class of evidence, is best understood by example. From the practitioner's perspective, psychological character evidence's most immediately recognizable use is in proving or disproving criminal acts and mental states.

1. Criminal Acts.—In *State v. Hickman*,\(^{25}\) the defendant was charged with felony murder based upon a death that occurred in the course of forcible sexual abuse.\(^{26}\) At trial, Hickman testified that his sexual contact with the victim had been consensual, not forcible.\(^{27}\) In rebuttal, the prosecution called a psychiatrist, who testified that he had done a study of rapists, and Hickman fit the class of "hatred rapists."\(^{28}\) Because hatred rapists are likely to engage in forcible rather than consensual sex, according to this argument, the defendant probably committed the forcible sexual act with which he was charged.\(^{29}\) The Supreme Court of Iowa held that this was proper rebuttal.\(^{30}\)

Expert testimony can also help disprove a criminal act without a direct opinion being offered on the likelihood that the defendant committed the crime. One of the more creative examples of this arose in *United States v. Roark*.\(^{31}\) There, the primary evidence that the defendant was a co-conspirator in a bank robbery was her confession.\(^{32}\) The United States Court of Appeals for the Eleventh Cir-

\(^{24}\) *Id.*
\(^{25}\) *Id.*, 337 N.W.2d 512 (Iowa 1983).
\(^{26}\) *Id.* at 514.
\(^{27}\) *Id.* at 516.
\(^{28}\) *Id.*

29. The traditional amount of force necessary to commit forcible rape is force sufficient to overcome the woman's resistance. See *Sue Bessem*, *The Laws of Rape* 59 (Landmark Dissertations in Women's Studies Series, Anne K. Baxter, ed., 1984). "Force" is thus logically relevant to proving lack of consent, an element of rape. See *id.* at 59-60. The focus in *Hickman*, however, was not on the victim's mental state (her "consent") but on the acts of the defendant: did he or did he not use violence to achieve sexual intercourse? See *337 N.W.2d* at 513-14, 516.

30. *Hickman*, 337 N.W.2d at 516. It is important to note that the defendant in *Hickman* had not offered character evidence, lay or expert. Rather, Hickman simply took the stand and testified that the sexual intercourse was consensual. *Hickman*, 337 N.W.2d at 516. It is, therefore, questionable whether under Rule 404 and considerations of undue prejudice the court was correct in permitting the prosecution to make the first use of expert character evidence on the grounds that it was proper "rebuttal." See *Fed. R. Evid.* 403, 404(a); *infra* notes 245-250 and accompanying text. Nor did the *Hickman* court examine the science underlying the prosecution expert's testimony or whether values other than truth-seeking justified admitting it. See *infra* Parts III.B.4 & III.C, notes 664-667 and accompanying text.

31. 753 F.2d 991 (11th Cir. 1985).

32. *Id.* at 992.
cuit held that the trial court erred in excluding the testimony of a defense psychiatrist who would have testified that Roark confessed only because of her “highly suggestible personality”—that she made up the confession to please her examiners. The appellate court held that the expert’s testimony would clearly assist the trier of fact in determining whether it was “more or less probable that Roark was somehow psychologically coerced into making the inculpatory statements.” If the confession were untrue and there was no other convincing evidence that Roark had conspired with the bank robber, she should not have been convicted of that crime.

2. Guilty Minds (the Susceptibility Defenses).—When psychological character testimony is offered to prove guilty minds, a common theme runs through many of the cases, a theme characterized in this Article by the term “susceptibility defenses.” This term encompasses the most common defense argument in favor of the admission of psychological character evidence to prove the defendant’s state of mind: Because of some unusual trait, the defendant is more or less susceptible than most to an extreme reaction to a particular type of situation. Thus, a defendant might be more fearful, more impulsive, more gullible, or more easily tempted or coerced than others. Psychological character testimony offers a “plausible alternative to normal” that asks the jury to take the defendant’s unusual weaknesses into account.

a. Negating Mental State.—One of the primary uses of psychological character testimony to bolster a susceptibility defense is to negate entirely a mental state required as an element of a crime. For example, a defendant may argue that when he picked up a gun upon the arrival of FBI agents to arrest him on another charge, he did so because he contemplated harming himself, not the agent. A psychologist’s testimony that the defendant was the type of person more likely to turn his aggression toward himself than others would thus be relevant and helpful against a prosecution for assaulting a federal officer. Similarly, a defendant may contend that his compulsion to gamble, as demonstrated by his diagnosed “pathological gambling disorder,” would make it unlikely that his failure to

33. Id. at 994-95.
34. Id.
36. See United States v. Staggs, 553 F.2d 1073, 1073-76 (7th Cir. 1977).
37. See id.
pay taxes was "wilful," because his financial mismanagement came from a need to gamble, not from a knowing effort to evade taxes.\textsuperscript{38}

\textit{b. Mitigation.---}A second use of expert character testimony for a "susceptibility defense" arises when a defendant argues that she lacked the mental state for the most serious crime but may still be guilty of a lesser offense.\textsuperscript{39} One critical application of this approach is when a defendant argues that because of some personality trait, such as impulsiveness under stress, an alleged murder lacked premeditation. This finding would mitigate the crime from first-degree to second-degree murder. A second application is when it can be demonstrated that the defendant acted in the heat of passion—and did so reasonably—which would reduce the crime to manslaughter.\textsuperscript{40}

An analogous effort to negate "specific intent" was made in \textit{United States v. Roberts},\textsuperscript{41} in which the defendant was charged with possession of cocaine with the intent to distribute.\textsuperscript{42} Roberts testified that he was a vigilante engaged in a one-man undercover operation to catch a major drug dealer and that he therefore possessed the cocaine not with the intent to distribute, but merely as a way of

\begin{footnotes}
\textsuperscript{38} See \textit{United States v. Shorter}, 809 F.2d 54, 58 (D.C. Cir.), \textit{cert. denied}, 484 U.S. 817 (1987). Neither the parties nor the court in \textit{Shorter} viewed the proffered expert testimony as "character" evidence. Nevertheless, that is essentially what was involved, because the defense expert was called to explain that Shorter's "cash lifestyle" was characteristic of compulsive gamblers as a group. See \textit{id.} at 59; see also infra text accompanying notes 119-135 (defining and illustrating "group" character evidence). This testimony was offered to respond to the government's claim that Shorter's lifestyle was evidence of his efforts to evade taxes. See \textit{Shorter}, 809 F.2d at 59. The defense apparently viewed that lifestyle as an aspect of obsession with gambling, not an indication of an effort to avoid paying taxes. See \textit{id.} In any event, the appellate court affirmed the exclusion of the expert testimony because it failed the \textit{Frye} test and was, in the court's view, unhelpful to the jury. See \textit{id.} at 58-62. For a discussion of the \textit{Frye} test, see infra text accompanying notes 266-346.

\textsuperscript{39} Obviously, "negating mental state" and "mitigation" both involve disproving a particular mental state. However, the former is a complete defense while the latter merely reduces the severity of the crime of which the defendant is convicted. This distinction is deemed important in substantive criminal law, see, e.g., Dressler, \textit{supra} note 4, at 474 (discussing a "mitigating defense" to homicide), and is therefore treated as important here.

\textsuperscript{40} See, e.g., \textit{State v. Christensen}, 628 P.2d 580, 583 (Ariz. 1981) (suggesting that psychiatrist's opinion that defendant had character trait of "acting without reflection" could mitigate first-degree to second-degree murder or even manslaughter); \textit{State v. Hallman}, 668 P.2d 874, 878 (Ariz. 1983) (holding that expert testimony regarding defendant's "impulsive personality" was relevant to negate mental state for first-degree murder).

\textsuperscript{41} 887 F.2d 534 (5th Cir. 1989).

\textsuperscript{42} \textit{Id.} at 535.
\end{footnotes}
catching a wrongdoer. The United States Court of Appeals for the Fifth Circuit held that the testimony of Roberts's psychiatrist—which illustrated the defendant's naive and autocratic personality—was consistent with the defendant's claims and was both relevant and admissible.

c. Affirmative Defenses.—A third and final major function of psychological character testimony, as used to bolster a special susceptibility claim, is to establish the elements of an affirmative defense. Different susceptibilities may be relevant to different defenses. For example, the susceptibility to extreme fear of physical harm in certain situations helps to establish self-defense in jurisdictions that take a subjective approach to this type of justification. In such jurisdictions, a defendant's honest belief that deadly force was necessary to prevent the defendant's imminent death or serious bodily injury is a complete defense to homicide or assault. It is therefore irrelevant that the defendant may not have behaved reasonably if the jury is persuaded that, because of the defendant's unusual susceptibility, she was honestly in fear. The battered woman syndrome is the best-known example of an argument of susceptibility to this type of fear.

Even in jurisdictions where self-defense claims may succeed only if the defendant acted as would a reasonable person under the circumstances—and the defendant's unusual traits might not be deemed an appropriate part of the reasonable person's "circum-

43. See id. Presumably (although not discussed by the court) the defendant would be guilty of simple possession, even under his version of the facts; it is for this reason that this case is offered as an example of mitigation, instead of a complete defense.

44. See id. at 536. Despite so holding, the court affirmed the conviction, finding the error harmless, given the "very strong evidence to support Roberts' conviction ..." Id. at 537.

45. See WAYNE R. LAFAVE & AUSTIN W. SCOTT, JR., CRIMINAL LAW 454-58 & n.28 (2d ed. 1986). The subjective approach is also followed by the Model Penal Code for crimes requiring more than negligence or recklessness. See MODEL PENAL CODE §§ 3.04(1) explanatory note, 3.09(2),(3) (1985).

46. See, e.g., State v. Hodges, 716 P.2d 563, 569 (Kan. 1986) (explaining that battered woman syndrome testimony is relevant to a self-defense claim, both in jurisdictions applying a subjective test and those applying an objective one). Such testimony is relevant even in objective jurisdictions because it must still be shown that the defendant's reasonable fear was honestly held. See Commonwealth v. Light, 326 A.2d 288, 294-95 (Pa. 1974) (Manderino, J., dissenting) (providing a particularly well-reasoned explanation of why expert testimony relating to subjective mental state is relevant in jurisdictions taking an objective approach to self-defense).

stances"—a defendant may still use evidence of unusual susceptibilities to support an "imperfect self defense" claim. To do so, the defendant must show that, although a complete defense is barred because the defendant's beliefs were unreasonable, the crime should be reduced from murder to manslaughter because the defendant's fears were nevertheless honestly held.49

Another type of susceptibility used to establish an affirmative defense is based on a defendant's susceptibility to the threats of others. While this appeal may involve fear as an element, its emphasis is on a defendant's weak will, her unwillingness to resist. The most obvious example involves the defense of duress or coercion.50 A defendant who was "passive and easily led" could therefore offer a psychologist's testimony to support a claim that she was an "unwilling participant" forced by a cohort into participating in theft, daytime housebreaking, and felony murder.51 Likewise, a defendant could argue that expert testimony regarding his relationship with his father would demonstrate that he was especially susceptible to, and was indeed, coerced by his father into assisting in murder.52

It is worth mentioning one further type of susceptibility that can lead to an affirmative defense: susceptibility to the influence, urgings, and persuasion of others. Here the emphasis is on gullibility, temptation, and reaction to others' persuasive powers, not necessarily passivity or fear of physical harm. This susceptibility is often most relevant to a subjective entrapment defense, which generally requires the defendant to offer some evidence of inducement by

48. See id. at 413-22 (discussing whether the "reasonable person" standard should include a battered woman's emotional and intellectual characteristics).

49. DRESSLER, supra note 4, at 199 (defining imperfect self-defense); see also Simmons v. State, 313 Md. 33, 38-45, 542 A.2d 1258, 1260-64 (1988) (explaining that expert testimony should be admissible to support imperfect self-defense when the testimony suggests that the defendant's psychological profile was consistent with an honest belief that deadly force was necessary).

50. "Duress" or "coercion" is a complete defense to many crimes where:
[a] person's unlawful threat (1) which causes the defendant reasonably to believe that the only way to avoid imminent death or serious bodily injury to himself or to another is to engage in conduct which violates the literal terms of the criminal law, and (2) which causes the defendant to engage in that conduct . . . .
LAFAVE & SCOTT, supra note 45, at 432-33. The defense is ordinarily not available in homicide prosecutions. See DRESSLER, supra note 4, at 259. But see infra text accompanying notes 51-52 (discussing homicide cases in which defense was raised).


52. See State v. Anderson, 379 N.W.2d 70, 78-79 (Minn. 1985), cert. denied, 476 U.S. 1141 (1986). The defendant's arguments were, however, ultimately unsuccessful. The appellate court affirmed the trial court's exclusion of psychiatric testimony regarding duress partly because it concerned, in the trial court's view, a matter within the common experience of jurors. See id. at 79.
government agents to commit the crime. If that burden is met, the onus shifts to the government to prove that the defendant was predisposed to commit the crime; absent such proof, the defendant prevails.\textsuperscript{53} Thus, a clinical psychologist's testimony that a defendant was particularly susceptible to persuasion and psychological pressure has been held admissible to prove that the government induced the defendant to distribute narcotics.\textsuperscript{54} Similar testimony also may be relevant to establish that a defendant was not "predisposed" to commit the crime, that is, not the type of person likely to commit the crime.\textsuperscript{55}

B. Criminal Justice and the Psychology of Moral Attribution

Having defined psychological character evidence and illustrated some of its practical uses at trial, we may now examine the major assumption in this discussion thus far, namely, that there is a deeply felt sense that a moral system of criminal prosecution should judge a defendant by that defendant's unique thoughts, feelings, and actions. This task will be accomplished by examining the concepts of "rule and case logics," developed by Professors Lempert and Sanders,\textsuperscript{56} which help to explain the interrelationship of the process of proof, the rules of criminal liability, and the social process of attributing moral responsibility.

1. Rule Logics.—Lempert and Sanders ground their arguments heavily in psychological research regarding when and why we hold one another accountable for violations of social norms. The re-

\begin{itemize}
\item \textsuperscript{54} United States v. Hill, 655 F.2d 512, 516 (3d Cir. 1981). \textit{Accord} United States v. Newman, 849 F.2d 156 (5th Cir. 1988) (explaining that testimony that a defendant had an intense reaction to perceived situational threats would have been admissible to prove inducement in an entrapment defense claim but for trial counsel's untimely, lengthy, and inconsistent offer of proof).
\item \textsuperscript{55} See, e.g., \textit{Hill}, 655 F.2d at 516. \textit{Accord} United States v. McLernon, 746 F.2d 1098, 1115 (6th Cir. 1984) (noting that "expert testimony concerning a defendant's predisposition may be invaluable in an entrapment case," but affirming the exclusion of such testimony because of inadequate notice).
\item \textsuperscript{56} See \textit{Richard Lempert & Joseph Sanders, An Invitation to Law and Social Science: Desert, Disputes, and Distribution} 27-30, 60-66 (1986). Lempert and Sanders's theories are presented as reflecting basic human needs. It may be, however, that cultural factors modify the analysis somewhat. For example, mental state is important in both the United States and Japan in making responsibility judgments. See V. Lee Hamilton \textit{& Joseph Sanders, Everyday Justice: Responsibility and the Individual in Japan and the United States} 127, 134 (1992). However, Americans place a heavier weight on mental state and find it significant in a wider variety of situations than do the Japanese. See \textit{id.} at 120-22, 124-34.
\end{itemize}
search reveals that moral responsibility judgments are guided by
two notions: (1) intention and (2) ability to do otherwise.57 Although Lempert and Sanders do not precisely define the word
"intention," they apparently mean something similar to the Model
Penal Code's definition of "purpose": acting with the awareness of
particular attendant circumstances and with the "conscious object"
of engaging in conduct of a particular nature, or of achieving a par-
ticular result.58 "Ability to do otherwise" encompasses a wide range
of concerns, including whether an actor (1) had the opportunity to
act differently, (2) had the literal physical ability to do so, (3) could
have learned to do so, and (4) would have had particular difficulty in
doing so.59 Legal concepts like constraint, compulsion, duress, mis-
take, unforeseeability, and irresistible impulse all embody notions of
inability to do otherwise.60

Although "full" moral responsibility requires both intention
and ability to do otherwise, we sometimes consider persons respon-
sible but not "fully so."61 Legal rules similarly often require less
than full moral responsibility before an actor may be held accounta-
ble.62 "Rule logics" define the degree of moral responsibility re-
quired by a particular rule of substantive law.63

Lempert and Sanders posit four types of rule logics: (1) the
criminal rule logic (for malum in se crimes), requiring both intention
and the ability to do otherwise; (2) contractual liability, requiring
only intention—to agree—but rendering ability to do otherwise ir-
relevant; (3) negligence liability, requiring only ability to do other-
wise but ignoring intention; and (4) strict liability, for which neither
intention nor ability to do otherwise is relevant.64

These four rule logics represent "ideal types" used for clarity of
exposition, but they oversimplify the real world.65 The real world of
law may contain rules demonstrating varying degrees of concern
with intention and ability to do otherwise beyond that embodied in
the four ideal types. For example, crimes requiring "recklessness"

58. MODEL PENAL CODE § 2.02(2)(a); see also LEMPERT & SANDERS, supra note 56, at
31-32 (distinguishing true intention—"meaning to do it"—from recklessness and mere
accident).
59. See LEMPERT & SANDERS, supra note 56, at 27 & n.7.
60. Id. at 33.
61. See id. at 27-28.
62. See MODEL PENAL CODE § 2.02(2)(c) (explaining that recklessness merely requires
a conscious disregard of a substantial and unjustifiable risk).
63. LEMPERT & SANDERS, supra note 56, at 27-30.
64. Id. at 28.
65. Id. at 3, 30, 35.
require proof that an actor was consciously aware of a substantial and unjustifiable risk of harm. Because actual awareness of the risk is required, recklessness embodies a closer concern with intention than does a negligence rule, but still falls short of true "intention" or desire to bring about harm. The real world is thus populated by spectrums of rule logics.

Rule logics, which specify what must be proved, are meaningless, however, without an understanding of how the requirements of substantive rules are proven. The process of proof is controlled by case logics.

2. Case Logics.—Case logics turn on the difference between "actor" and "social" meanings. An "actor meaning" is the meaning an actor actually gives to his own behavior. A "social meaning" is the normal or typical meaning of behavior. Observers often attribute social meanings to another individual's actions instead of truer actor meanings. This may be so for many reasons, but one of the most important is that observers are often unaware of "the variety of historical and contextual forces that impinge on [an individual actor's] behavior" in ways that are not immediately obvious. The likelihood that an observer will come to understand a particular actor's meaning—her true thoughts and feelings—is thus governed to a large extent by how much the observer learns about these forces. How much the observer learns is in turn controlled by the observer's willingness to recognize the "uniqueness" of the actor's situation and to engage in an "extensive" inquiry.

66. See Model Penal Code § 2.02(2)(c).
67. See id. § 2.02 cmts. 3-4 (defining recklessness); see also Lempert & Sanders, supra note 56, at 31 n.10 (justifying criminal recklessness as either demonstrating a depraved nature or serving as a proxy for true intent).
68. Lempert & Sanders, supra note 56, at 51-52.
69. Id. at 51.
70. Id.
71. Id. at 51-52.
72. "Uniqueness" is defined as "the degree to which we attend to the special characteristics of a situation in identifying its legal implications." Id. at 59. Its opposite, "typicality," focuses on broad descriptions of commonalities instead of on the details that make each action unique. See id. at 43-44, 51-52. Obviously, there are wide ranges in the degree of detail we observe in any particular situation and in our willingness to view certain types of events as unique.
73. Id. at 45-51. "Extensiveness" refers to "the degree to which we search out and accept evidence that is temporally or spatially removed from the event in question." Id. at 59. Its opposite, "immediacy," focuses solely on what happened at a particular time and place. See id. at 49-51, 62-63. Again, the real world involves a wide range of variation in the degree to which we inquire into the history leading up to a particular event.
The two dimensions of uniqueness and extensiveness are, of course, interrelated: focusing on what is unique in an individual may require one to learn more about that individual's history, while learning more about that history may make the individual's situation appear more relatively unique.\textsuperscript{74} The central point, however, is that when observers employ immediate, typical analyses—what Lempert and Sanders call "shallow case logics"\textsuperscript{75}—they are likely to ascribe social meanings to an actor, for example, to ascribe a "normal" mental state to the actor even though the actor's mental state may have been different from what a normalized account might suggest.\textsuperscript{76} This tendency is magnified when circumstances fit a convenient stereotype, for then the attributor is more likely than ever to overlook the possibility that a unique actor-meaning accounted for the behavior.\textsuperscript{77} By contrast, when observers use unique, extended inquiries—"deep case logics"—they are more likely to uncover true actor-meanings.\textsuperscript{78}

These "ideal types"—shallow and deep case logics—determine whether rule logics have any real meaning in a particular context. Thus, shallow case logics may involve typifications that distort or ignore intentions in a civil negligence case. But these logics do little harm, for most negligence rules by definition ignore actor intentions.\textsuperscript{79} A pure criminal rule logic, by contrast, requires proof that an actor acted with purpose,\textsuperscript{80} yet a shallow case logic may result in a jury attributing purpose to the actor when he had none, a truth that might have been revealed by a closer examination of the actor's history and of the special qualities that make him unique.\textsuperscript{81} A shallow case logic may thus lead to unreliable results, inconsistent with the rule logics that are the moral bases underlying the substantive rules of criminal liability.\textsuperscript{82}

\textsuperscript{74} Id. at 59.
\textsuperscript{75} Id. at 63.
\textsuperscript{76} See id. at 63-64 (comparing examples of actor- and social-meanings).
\textsuperscript{77} See id. at 63 ("When circumstances and behavior fit a convenient stereotype, an attributor is less likely to consider the possibility that a particular actor meaning accounts for behavior.").
\textsuperscript{78} Id.
\textsuperscript{79} See W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 32, at 173-74 (5th ed. 1984) (explaining that an action in negligence involves an objective inquiry to determine what a "reasonable person" would do, rather than a subjective approach to ascertain the defendant's intentions).
\textsuperscript{80} See LEMPERT & SANDERS, supra note 56, at 28.
\textsuperscript{81} See id. at 63-66.
\textsuperscript{82} Id. at 85, 131.
3. Typification and How to Overcome It.—Yet Lempert and Sanders report a strong natural tendency to employ shallow logics in both civil and criminal cases.\[83\] This tendency alone requires special efforts by judges and attorneys in criminal cases to move fact-finders toward the deeper analyses required by the substantive criminal law. Apart from this natural tendency toward shallow logics in every case, however, there are special forces at work in criminal cases that amplify the likelihood of fact-finders using typified, immediate analyses and thus make even more urgent the need to respond.

First, the limited resources available to the criminal justice system result in a burden of cases that create pressure for judges and lawyers to move cases quickly, devoting little time to any individual case.\[84\] This “assembly-line justice” becomes depersonalized and dehumanized, resulting in a reluctance to treat defendants as individuals; instead, the defendants are viewed stereotypically and their cases are dropped into neat, predefined categories, with little effort made to explore what may be special or different about particular defendants.\[85\]

Second, criminal cases often involve highly emotionally charged issues reflecting broader societal conflicts, values, and fears. Thus, deep-seated stereotypes and myths about racial groups, sex-roles, and cultural groups may be especially powerful, amplifying the likelihood of typified analyses.\[86\] When a criminal defendant has had unusual life experiences or is in some way emotionally dysfunctional—but not to the point of mental illness—jurors are even more likely to draw on “superstition, fear, and popular wisdom.”\[87\]

Third, the high stakes and sometimes troubling emotional issues raised in criminal cases may exaggerate the effects of limited information. An analogy to the economist’s explanation of racial discrimination may help to explain why this is so. Economists rely on the theory of “information costs,” which, as its title implies, emphasizes that there are costs associated with obtaining information.\[88\] It would thus be costly to seek out information contrary to

\[83\] Id. at 86.
\[84\] Id. at 63, 75-78.
\[85\] See Herbert L. Packer, The Limits of the Criminal Sanction 292 (1968) (“So great is the crush of cases and so little is the time available to deal with any given one that there is an almost total lack of attention to the situation of the individual defendant. The process is depersonalized and dehumanized.”).
\[86\] Lempert & Sanders, supra note 56, at 63; see infra text accompanying notes 119-118, 129-135.
\[87\] Bonnie & Slobogin, supra note 35, at 492.
the teachings of one's parents, community, or one's own experience. Racial stereotyping, for example, saves people from incurring these costs, because it allows people to make decisions about others based on generalizations, instead of forcing them to seek out specific information regarding the particular person in question. Over time, such stereotypes may be abandoned if costs are imposed on the racist for using the stereotype, for example, losing the best workers to those who are not racists.89

In a trial, however, there are severe constraints on the information that jurors may obtain, for the lawyers and the judge control the flow of information. Moreover, the trial ends after days or weeks, with no impact thereafter on the jurors' lives, so that time passage cannot offer incentives—like the loss of good workers—for the jurors to abandon their stereotypes.90 Finally, the jurors know that their decision may imprison someone, a heavy emotional burden that may require them to receive assurance from somewhere that their decisions about guilt or innocence are accurate ones. These are powerful pressures on individual jurors to resort to stereotyping for guidance.

The most effective weapon for combating each of these pressures toward shallow logics is information. That is precisely what psychological character evidence provides. Lempert and Sanders explain that an adequate informational base creates powerful "empathetic pressures" to understand actor-meanings.91 These pressures can help jurors overcome stereotypes and may even lead judges to switch mid-trial from their usual use of shallow logics and demand for speedy case-processing to the employment of genuinely unique, extended analyses that respect the criminal law's ethic of individualized justice.92 Indeed, trial lawyers are instinctively aware both of the dangers of typifications and of the value of psychological character evidence as a response.

One such trial lawyer, Susan B. Jordan, an attorney who has used psychological character testimony to support self-defense claims by women charged with killing their rapists and physical

89. Id. at 352-54, 362.
90. Cf. id. at 236-37 (noting that jurors incur no cost from behaving irrationally and thus should be screened from stale evidence regarding a criminal defendant's prior criminal acts).
91. LEMPERT & SANDERS, supra note 56, at 65.
92. See id.
abusers, described counsel's job in such cases as follows:

What generally happens in a criminal trial is that the District Attorney puts on trial one instant in time; did the defendant pull the trigger? From the DA's point of view, that is basically all he or she has to prove, unless he is trying to prove a case of premeditation, in which case he must show some planning in that. But he doesn't go much further back in time to explain how the defendant came to be in this spot.

But when someone comes to pull a trigger, there are significant social and psychological reality events which preceded that. No matter who the person is, something, often a very complex set of things, preceded the pulling of that trigger. In my opinion, the defense attorney's job is to explain to the jury what psychological and social factors put this person in the defendant's chair . . . .

C. Morality and the Principles of Criminal Evidence

Lempert and Sanders's theory of case logics implicitly suggests an important underlying principle: major differences exist between the rules of criminal and civil evidence because the former involves significant ethical precepts that the latter does not. This principle is a corollary of the notion that criminal liability is justified primarily by moral values, not economic or political ones, and that one of the major functions of criminal punishment is therefore to express society's moral condemnation. If the criminal law is to serve this function well, the substantive rules of criminal liability must require


94. Id. at 466-67; see also Brooks, supra note 11, at 23-24 (noting value of psychological testimony in countering jurors' normal assumptions).

95. See Lempert & Sanders, supra note 56, at 131 ("[W]hen criminal sanctions are at issue the adoption of a shallow case logic, which fails to give full consideration to questions of agency and purpose, threatens the moral foundation on which the law must rest."); see also A. Zuckerman, The Principles of Criminal Evidence 6 (1989) ("The law of criminal evidence has an independence of purpose and a cohesion of principle such that it requires a discrete treatment.").

96. See Packer, supra note 85, at 116 (discussing the prominent social purposes of the criminal law).

97. See Igor Primoratz, Justifying Legal Punishment 149-54 (1989) (discussing the value of moral condemnation); see also Kip Schlegel, Just Deserts for Corporate Criminals 72-74 (1990) ("Regardless of whether the purpose of punishment is primarily crime prevention or retribution, the fact is that punishment contains a condemnatory function.").
full moral responsibility, and the criminal law's notions of responsibility in turn must be generally consistent with popular notions of morality. While concerns about efficiency, the peaceful settlement of disputes, and the distribution of social justice, that is, access to wealth, education, and power, may justify a lesser concern with morality in the law of torts or contracts, and thus something significantly less than "full" moral responsibility, in the criminal law nothing less will do.

As noted earlier, full responsibility requires a concern with both ability to do otherwise and intentions. Indeed, a concern with intentions is demonstrated by the strong trend of the substantive criminal law in recent years toward increased subjectivity, that is, an increased commitment to mental states and actor-meanings.

This subjectivity reflects the criminal law's basic concern with "just deserts": the notion that a criminal should receive the punishment that he "deserves," a punishment that must therefore be proportionate to the crime. Proportionality depends on the degree of harm inflicted or risked and culpability. Culpability turns on mental state, motives, and excuses. Thus, greater punishment is deserved when the same harm proceeds from a more depraved mental state.

98. See LEMPERT & SANDERS, supra note 56, at 35-36, 38-39 (explaining the relationship between "popular morality" and "fuller-mind responsibility").
99. Id. at 39-40.
100. Id. at 134.
101. Id. at 283-84.
102. See id. at 35-41.
103. See supra notes 57-61 and accompanying text.
104. Bonnie & Slobogin, supra note 35, at 432 n.8. Some courts and commentators have argued that, at least when punishment and moral stigma are severe, the requirement of wrongful subjective mental states is constitutionally required. See, e.g., Commonwealth v. Heck, 491 A.2d 212, 218-28 (Pa. Super. 1985) ("[T]he legislature's power to eliminate mens rea is not without limitation, and the exercise of that power must comport with constitutional principles."); aff'd, 535 A.2d 575 (Pa. 1987); James J. Hippard, Sr., The Unconstitutionality of Criminal Liability Without Fault: An Argument for a Constitutional Doctrine of Mens Rea, 10 Hous. L. Rev. 1039 (1973) (arguing that strict liability in the criminal law is unconstitutional).
105. JEFFRIE G. MURPHY & JULES COLEMAN, PHILOSOPHY OF LAW: An INTRODUCTION to JURISPRUDENCE 120-24 (1990) (just deserts); PRIMORATZ, supra note 97, at 12 (proportionality).
106. SCHLEGEL, supra note 97, at 48-49.
107. Id.
108. See id. at 117, 129-33 ("[A defendant] merits descending levels of punishment (holding harm constant) by virtue of the reduced blame-worthiness associated with the act."); accord MURPHY & COLEMAN, supra note 105, at 128 ("Justice demands individuation . . . . And it seems elemental that there is a significant moral difference between a person who kills maliciously and one who kills accidentally. . . . ").
Similarly, many full or partial excuses, such as provocation and heat of passion, require a subjective inquiry into the mind of the accused. 109

State power, specifically the power to imprison and to condemn offenders to society's most extreme moral condemnation and loss of standing in their local communities sets the law of criminal evidence apart from that of civil evidence. 110 The fear of state power, combined with notions of just deserts and the moral educative role of the criminal law, requires that the procedures by which criminal liability is imposed be perceived as both fair and accurate and that the risk of convicting the innocent be minimized. 111 This latter concern is embodied in the constitutional requirement that guilt be imposed only if the state meets its burden of establishing the elements of a

109. See Primoratz, supra note 97, at 91-92 (discussing moral bases for mitigation); Dressler, supra note 4, §§ 31.04[c][4], 31.08[A][C][1]-[2] (discussing subjective components of heat of passion and provocation).

There is no need to resolve here the ongoing debate between utilitarians—those who judge the morality of punishment by its future good consequences, see Primoratz, supra note 97, at 12—and retributivists—those who judge the morality of punishment by the past injustice committed, see id. Nor could such a task sensibly be attempted in this Article, for most now agree that there is no inherent inconsistency between utilitarianism and a concern with proportionate punishments.

Where the respect-based theory of retribution will always have a valuable role, however, is as an absolute side-constraint on the pursuit of utilitarian deterrence. As such it will remind us that, in our zeal to pursue the social safety that we hope will flow from an effective system of deterrence, we must not be willing to buy this benefit by treating criminals more harshly than they deserve to be treated.

Murphy & Coleman, supra note 105, at 124. Indeed, utilitarian justifications for the notion of "just deserts" are both possible and logical. See Packer, supra note 85, at 16, 112-13 (emphasizing utilitarianism's need to balance its traditional concern with crime deterrence against supposedly "retributive" concerns with impact on human autonomy and the opportunity for free choice); Schlegel, supra note 97, at 66-67, 74 (arguing that concept of "just deserts" can be grounded in both utilitarian and retributive concerns). But cf. Primoratz, supra note 97, at 111-43 (criticizing theories which attempt to incorporate both utilitarianism and retributivism). At the very least, most recognize that these subjective, retributive concerns with mental states, mercy, and excuses serve as useful and necessary restraints on state power. See Murphy & Coleman, supra note 105, at 124; Packer, supra note 74, at 112-13 ("[C]ulpability is an appropriate criterion for limiting the reach of state intervention."); see also Jeffrey Sedgwick, Law Enforcement Planning: The Limits of an Economic Analysis 162-63 (1984) (arguing that even the economic approach to law must recognize that moral, retributive concerns with right and wrong promote deterrence). But see the argument of Richard Posner, An Economic Theory of the Criminal Law, 85 Col. L. Rev. 1193, 1195, 1230-31 (1985), who concludes that economic efficiency, not morality, is the primary objective of the criminal law, a view soundly debunked in Jules L. Coleman, Markets, Morals and the Law 164-65 (1988).

110. Zuckerman, supra note 95, at 4-6.

111. See id. at 5 ("[N]ot only must the criminal process protect the innocent but it must also be seen to be doing so.").
Unfortunately, the growing formal commitment of the substantive criminal law to notions of just deserts and subjective actor-meanings—in short, to "individualized justice"—has been checked by a bureaucractized system of prosecution and a focus on crime control over procedural fairness or minimization of the risk of error. When the law of criminal evidence fails to move fact-finders toward deep case logics that maximize the likelihood of outcomes consistent with theories of just deserts, the moral basis of the criminal law is undermined. Lempert and Sanders summarized this danger well:

[What is . . . disturbing to many is the apparent ease with which lawyers and judges adopt a shallow case logic in the case of accused criminals. Shallow case logics restrict the law's commitment to the full humanity of individual actions. It is true that shallow case logics tend to reduce transaction costs in the criminal area as they do with torts, but the savings do not have the same implications, for the criminal law is not primarily concerned with the allocation of costs. Bureaucratic adjudication, as Weber feared, undermines the ethical dimension of law embodied in the concept of desert. In situations where individual desert is the central concern, we risk important values when we allow adjudications to be transformed into a mode of bureaucratic decision making. One solution is to refocus our concerns and forego moral judgments. With the spread of insurance and the extension of strict liability, this has been happening to a large extent in the tort law. If, on the other hand, we remain interested in judging the morality of behavior, and this is what the criminal law purports to do, we are, if we take the task seriously, committed to a deep case logic.]

What is the consequence, however, of this commitment to a deep case logic? Such a commitment certainly cannot require the
admission of all evidence that tends to promote unique and extended inquiries. Much evidence of that nature may indeed mislead or confuse the jury, undermining a fair, open-minded focus on actor-meanings. In any event, there are competing values, other than individualized justice, to be served by the law of evidence. However, the tendency toward typification in criminal cases does require trial judges to be particularly alert to the impact that evidentiary rulings may have on case logics. The need to promote deeper logics must be an important value, a moral value, that informs the judge's weighing process.

Notably, the need for deep logics is not always in the defendant's benefit, for such an inquiry may at times reveal actor-meanings that support, rather than defeat, his guilt. While the occasional prosecutorial efforts to deepen case logics are also entitled to an attentive ear, the constitutional requirement of proof beyond a reasonable doubt embodies a judgment that the risk of error regarding guilt should fall on the prosecution, not the defendant. In deciding whether to admit or to exclude evidence that may impact upon case logics, therefore, a judge should resolve her doubts about how to weigh the impact against other concerns in favor of the defense.

How the courts can adequately consider the role of case logics and the moral imperatives of the law of criminal evidence in resolving questions concerning the admissibility of psychological character evidence is the subject of much of Part III of this Article. For now, however, this Article turns to a more modest goal: illustrating how psychological character evidence can help move juries closer to the individualized justice that social psychology, moral principles, and the rules of the substantive criminal law demand.

D. Individualized Justice and the Spectrum of Typification

The theoretical dichotomy suggested thus far between stereotypical, "normalized," justice and "individualized" justice is to some extent a false one. In the "real world" there are many levels of typification, each varying in the degree to which character evidence seeks to portray the defendant as unique, or special, rather than typical. All character evidence involves at least an implicit com-

116. See infra text accompanying notes 330-344, 495-497.
117. See LEMPERT & SANDERS, supra note 56, at 65 ("Some actions appear less culpable when subjected to closer scrutiny, while others may appear more culpable.").
parison to a group claimed to be similar to, or different from, the defendant, so a pure focus on the individual does not reflect reality. Nevertheless, the degree to which the group is emphasized over the individual may vary.

Group or individual character evidence, the polar opposites of this emphasis, are two overly simplistic but useful categories for understanding the ways in which psychological character evidence can help to achieve more individualized justice. Heavy emphasis on the group leads to a portrayal of the defendant in a relatively stereotypical fashion: a "typical" member of a certain group. On the other hand, heavy emphasis on the individual leads to the defendant being viewed as relatively unique, her membership in a group defined more by her own life experiences, her personal story. Thus, while these opposites are often utilized, it must be remembered that there is conceivably an infinite variation in the degree to which one is emphasized over the other in practice—that is, there is a spectrum of typification.

1. The Upper End of the Spectrum: The Group.—"Group character evidence" is a handy label for testimony in which an explicit comparison is made between the personality traits of an individual and those typical of members of a certain group. The testimony may be general, the expert merely describing the characteristics of the group and leaving the individual's membership in that group as a question for the jury. Alternatively, the testimony may be specific, with the expert testifying that the defendant displays those qualities necessary for membership in the group.

119. Glen Weissenberger, Making Sense of Extrinsic Act Evidence: Federal Rule of Evidence 404(b), 70 Iowa L. Rev. 579, 613 n.93 (1985) ("[T]o have meaning, the concept of propensity must be dependent upon our collective assumptions as to the repetition of certain types of human behavior based on the experience of observing persons other than the accused." (emphasis added)). Cf. David McCord, Syndromes, Profiles and Other Mental Exotica: A New Approach to the Admissibility of Nontraditional Psychological Evidence in Criminal Cases, 66 Ore. L. Rev. 19, 27 (1987) [hereinafter McCord, Profiles] ("All nontraditional psychological evidence follows this pattern of a comparison of an individual's behavior with a database consisting of the behavior of others in similar circumstances who have been studied in the past.").

120. But members of one group may have different qualities from those of another group or from an "average" person. By focusing on the special qualities of a defendant's particular group, instead of on preconceived or "normalized" notions of the defendant's likely qualities, group character evidence helps to individualize justice.

121. See McCord, Profiles, supra note 119, at 52 ("A moment's reflection on these categories of evidence reveals that 'group' character evidence is objectionable for the same reason as is traditional character evidence: probative value depends upon the jury drawing the forbidden inference that the defendant has a propensity to commit the crime with which he is charged.").
Group character evidence is often used in one of two ways: (1) in a purely probabilistic, predictive fashion or (2) to educate jurors or correct their misconceptions about a group.

a. Probabilistic Prediction.—An example of a probabilistic, predictive use of character evidence is when a defendant argues that he does not fit the personality "profile" of most rapists and, therefore, it is less likely that he committed that crime. A corresponding prosecution use is the claim that the defendant does indeed fit the profile, thus raising the likelihood of guilt. Note that both uses appeal to stereotypes, either pre-existing ones (e.g., the "normal male") or ones sought to be newly created in the jurors' minds (e.g., the "typical" rapist). Moreover, the primary function of this stereotyping is to prove that the defendant did or did not commit an act rather than the presence or absence of mens rea.

Probabilistic use raises some obvious concerns. For example, suppose that a psychiatrist testified that eighty-five percent of all child sex abusers are passive-dependent personalities, that figure being offered to prove that the defendant, who is also a passive-dependent, is guilty. This figure is of little value without a comparison to the total population of passive-dependents. If, therefore, such a comparison revealed that 99.9 percent of passive-dependent personalities are not child sex abusers, the fact that the defendant had such a personality would not suggest a propensity for child abuse.

Moreover, explicit probabilistic evidence raises concerns about overwhelming the jury with numbers so that it convicts based upon


123. Character evidence presented by the prosecution is usually done so on rebuttal, given the common ban on prosecution's first use of character evidence to prove conduct. See Fed. R. Evid. 404; infra text accompanying notes 244-265.


125. This hypothetical is based on the facts of In re Cheryl H, 153 Cal. App. 3d 1098, 1124 & n.28 (1984), a civil case, but one containing a useful critique of this abuse of probabilistic character evidence.
some perceived notion of the statistical likelihood of guilt instead of the evidence of what happened on the specific occasion in question.\textsuperscript{126} These concerns about absurd statistics and overawed juries have either explicitly or implicitly led some courts to reject the probabilistic use of character evidence as unhelpful to the jury's inquiry.\textsuperscript{127} Additionally, concerns that the jury will believe that predictions of future human behavior are infallible have led some courts to reject probabilistic character evidence absent proof of its "general acceptance" in the relevant scientific community or other proof of its predictive value.\textsuperscript{128}

b. Correcting Misconceptions.—The second use of group character evidence—to educate jurors or correct their misconceptions about some group—does not necessarily raise the same concerns as probabilistic character evidence. For example, a battered child claiming to have killed his brutal father in self-defense because he feared for his life might offer evidence that all battered children display extreme fear in certain situations. It should not matter that other, non-battered children might display the same symptoms. The point is that \textit{this child} displays those symptoms and, therefore, his story of acting in fear for his life is more plausible. Nevertheless, some courts have remained skeptical about the value of such evidence in demonstrating fear at the time of the crime absent lay testimony of other events supporting such fear.\textsuperscript{129}

Still other courts have found expert testimony of this nature unhelpful because lay testimony, for example about an extensive history of prior beatings of a woman or a child, is sufficient alone to support the defendant's claim of fear caused by the earlier beatings.\textsuperscript{130} These courts fail to appreciate that "extensive" testimony providing history and context may be insufficient without expert ad-


\textsuperscript{127} \textit{E.g.}, State v. Claflin, 690 P.2d 1186, 1190 (Wash. App. 1984) (prosecution expert's testimony on rebuttal in rape case that 43\% of child molestation cases were reported to have been committed by father figures was unduly prejudicial).

\textsuperscript{128} \textit{E.g.}, State v. Cavallo, 443 A.2d 1020, 1026 (N.J. 1982).


vice to demonstrate the "uniqueness" of the defendant's situation. Thus, testimony regarding the "cycle of violence" and "learned helplessness" of battered women (1) may add greater plausibility to the woman's testimony regarding the perceived degree of the beatings and the predictability of her husband's rage, (2) may explain why a woman might not leave her husband despite such beatings, and (3) may reveal why a woman might believe that her husband was about to kill or maim her, although the basis of her belief may not be obvious to others. The history of the beatings is simply not enough.

Another difference between the probabilistic use of expert character testimony and its use as a "corrective" or "information-enhancing" device, is that the former usually concerns acts while the latter concerns mental states. Whether this is a logical distinction regarding the value of the evidence or not, some courts have suggested that understanding another's mind rather than predicting another's behavior is more properly within a psychologist's expertise and traditional role in the courtroom. These courts are more likely to view character testimony offered to prove state of mind as already within the sphere of "generally accepted" psychological techniques. On the other hand, some courts continue to apply the traditional screening tests for scientific evidence to all psychological character testimony, even as to state of mind, examining the scientific bases with somewhat greater scrutiny.

One other form of the corrective or educational use of psychological character evidence involves the comparison of an individual to "normal," not for probabilistic predictive reasons, but to help the jury understand how normal people react in certain situations. For example, psychological character testimony may be offered to support the argument that the defendant's actions were consistent with those of a "reasonable" person. This use has generally been rejected on the theory that jurors already know what is normal and the

131. See Stonehouse, 555 A.2d at 783-85 (describing the use of battered woman syndrome in dispelling erroneous myths); Kinports, supra note 47, at 396-408 (detailing the cycle of violence, learned helplessness, and other features of the battered woman syndrome); Lenore Walker, The Battered Woman 18-31 (1979) (cataloguing the various myths regarding battered women).

132. See McCord, Profiles, supra note 119, at 35-58 (summarizing cases).


testimony is, therefore, unhelpful. This approach can be too simplistic in particular instances. For example, jurors may not be aware of the impact that a combination of particular environmental factors or "stressors" may have on the degree and nature of a normal person's fear. Thus, psychological expert testimony regarding these stressors might be admitted to support a defendant's claim that she killed because of an honest fear that she had to do so to protect a loved one.

2. The Lower End of the Spectrum: The Individual.—With "individual character evidence," unlike "group character evidence," the comparison to a group is either de-emphasized or merely implicit. The focus is instead on the individual's own history and unique personality. Less reliance is placed on comparison to systematic, statistical studies and experiments. Rather, emphasis is placed on an admittedly subjective, clinical judgment by the psychologist based upon personality tests, clinical experience with other patients, and interviews with the patient in question. This approach does not necessarily eliminate statistical references or explicit group comparisons, but such comparisons are not emphasized and are not the primary basis for the expert's opinion. Courts sometimes see these differences as important.

For example, in People v. Stoll, a child sexual abuse case, the trial court excluded a psychologist's opinion that the defendant had a normal personality function, with a low indication of antisocial or aggressive behavior, thus making it unlikely that he would commit the charged acts. The opinion was based on a combination of personality tests, interviews with the defendant, and the psychologist's professional experience. Moreover, the psychologist testified in court as to the statistical accuracy of the personality tests used,

135. E.g., Cynthia K. Gillespie, Justifiable Homicide 166 (1989) (noting judges sometimes exclude battered woman syndrome testimony because it is about "just ordinary fear—specifically the reasonableness of the woman's fear—and fear is a common human emotion that any juror is familiar with and can understand without help from an expert"); Neil J. Vidmar & Regina A. Schuller, Juries and Expert Evidence: Social Framework Testimony, 52 Law & Contemp. Probs. 133, 139 (Autumn 1989) ("A frequently expressed objection to social framework evidence is that it contributes little or no new information about human behavior to the pooled experiences of the jurors."). See also Braley v. State, 741 P.2d 1061, 1063-64 (Wyo. 1987). Justice Urbigkit, in a persuasive dissent, argued, however, that the psychology of conflict and fear is a complex subject on which expert testimony may be especially helpful. Id. at 1075-76 (Urbigkit, J., dissenting).
136. 783 P.2d 698 (Cal. 1989).
137. Id. at 703-06.
but noted that the tests were not the sole determinant of his opinion.\textsuperscript{138} On appeal, the State argued unsuccessfully that the testimony should be admitted only if valid empirical evidence supported both the existence of a rapist profile and that it was significant that a defendant did not fit that profile.\textsuperscript{139} In reversing, the appellate court refused to require a probabilistic comparison to a profile, holding the evidence admissible because the subjective clinical judgment of the psychologist was both relevant and helpful to the jury.\textsuperscript{140} The court further held that there was no need to subject the testimony to the kinds of special tests often applied to screen "scientific" evidence.\textsuperscript{141}

\textit{Stoll} is representative of "individual character evidence" in its purest form because the court rejected the notion that express comparison to a group is necessary. Instead, the court approved psychological testimony that was unquestionably based on the expert's subjective judgments and that sought to depict the defendant as a truly unique individual, a "universe of one." If such testimony proved successful on retrial, it would surely have achieved the goal of overcoming juror tendencies toward trial by typification, myth, and stereotype.

\section*{II. Tools of the Trade}

Part I of this Article reviewed the moral and psychological roles that may be played by expert character witnesses in criminal trials. It also illustrated the wide variety of potential uses for such witnesses in proving the elements of various crimes and defenses and briefly examined the value placed by the judiciary on such witnesses at trial. Before turning to a more detailed inquiry into the evidentiary objections to such testimony, however, it is necessary to understand what expert character witnesses do, for only then can it be judged when, if ever, they may be helpful to a jury.

Part II therefore offers a brief general review of the "tools" of the psychologist's trade. These tools include: objective and projective personality tests; clinical interviews; behavioral observations; reviews of empirical studies of personality types, syndromes, and profiles; and interviews with third parties. The tools reviewed are those common to a wide variety of character testimony. No effort is

\textsuperscript{138} Id. at 705.
\textsuperscript{139} Id. at 707.
\textsuperscript{140} Id. at 714-15.
\textsuperscript{141} Id. at 714.
made to examine any single tool in depth, although selected types may be emphasized for illustration.

These tools can be understood only, however, by first examining what psychologists mean when they speak of "personality," their equivalent to the legal concept of "character."

**A. The Concept of Personality**

"Personality" is the central psychological concept concerning the "type of person" one is. Although different theorists define the term in slightly different ways, one useful common definition is as follows: "Personality is a stable set of tendencies and characteristics that determine those commonalities and differences in people's psychological behavior (thoughts, feelings, and actions) that have continuity in time and that may not be easily understood as the sole result of the social and biological pressures of the moment."

This definition requires several explanations and qualifications. First, although the definition includes both cognitive components—those concerning how information is perceived, retained, transformed, and acted on—and affective ones—emotions, temperament, and the like—the traditional focus of personality study often emphasizes the affective over the cognitive. But the more modern trend—and the one followed in this Article—has been to include certain cognitive components in the study of personality.

Second, while some theorists include intelligence as a character trait, it is often treated as a separate subject of study. Moreover, courts often describe intelligence as *not* an aspect of character, at least when it is argued that, because of low intelligence, the accused lacked the capacity to form a mental state.

143. *E.g.*, Lewis R. Aiken, Assessment of Personality 4 (1989) [hereinafter Aiken, Assessment] (defining personality as including both cognitive and affective components but then focusing "on the assessment of affective characteristics, which encompass the traditional, albeit somewhat limited, conception of personality variables").
144. *See, e.g.*, B.R. Hergenhahn, An Introduction to Theories of Personality 309-40, 400-26 (3d ed. 1990) (surveying selected cognitive theories of personality); Maddi, supra note 142, at 155-83 (grouping certain cognitive personality theories as constituting the "consistency model" of personality theory).
145. *See* Aiken, Assessment, supra note 143, at 4.
146. *See* Maddi, supra note 142, at 263 (cautioning against reliance on a particular study because its "major concern is with intelligence rather than personality").
147. *E.g.*, United States v. West, 670 F.2d 675, 682 (7th Cir.) (concluding that intelligence is not a "character trait" within the meaning of Federal Rule of Evidence 404), cert. denied, 457 U.S. 1124 (1982).
capacity defense, not character evidence in the sense used in this Article.\textsuperscript{148}

Third, personalities can and do change,\textsuperscript{149} although theorists disagree over the likely speed and magnitude of such change.\textsuperscript{150} Understanding this dynamic nature of personality is an important insight. Because personality traits are not forever fixed, life events may cause personality change. Thus rape victims, victims of child abuse, and battered women may come to display behaviors, thoughts, and feelings that persist over some significant period of time, and thus may be considered part of their personality.\textsuperscript{151} So viewed, many of the psychological syndromes and profiles that have been the focus of much recent legal literature can be seen as aspects of a single problem, that of using expert psychological character evidence.

A fuller understanding of personality also requires an examination of the recent assaults in the academic community on the very concept of “personality.” The best-known assault comes from social learning theorist Walter Mischel, whose early research suggested that behavior has little consistency across time and situations, and that the notion of a “stable” set of behavioral tendencies—a personality—thus makes little sense.\textsuperscript{152} Indeed, some recent legal literature has picked up on Mischel’s research as a ground for challenging the value of character evidence generally.\textsuperscript{153}

\begin{flushright}
\textsuperscript{148} See infra text accompanying notes 575-584.
\textsuperscript{149} HERGENHAHN, supra note 144, at 178 (“Personality, according to Allport, is never something that is; rather, it is something that is \textit{becoming}. Although enough similarity exists within people to maintain their identity from one experience to another, in a sense they never are quite the same people they were before a particular experience.”).
\textsuperscript{150} MADDI, supra note 142, at 258-59. See also HERGENHAHN, supra note 144, at 3-6 (discussing generally the theories considering influences that make up personality).
\textsuperscript{151} The authors of one of the leading texts on law and social science have classified psychological testimony regarding “personality” as distinct from such testimony regarding “life events.” See JOHN MONAHAN & LAURENS WALKER, SOCIAL SCIENCE IN LAW: CASES AND MATERIALS 359, 366 (2d ed. 1990). The distinction is an unhelpful one when life events are offered primarily to explain personality changes. Both “personality” and “life events” testimony raise similar psychological issues and both involve the legal concept of “character” evidence. Indeed, the authors expressly recognize the latter point as to “personality” testimony, \textit{id.} at 365, and implicitly do so as to “life events” testimony by selecting cases addressing the “character evidence” question. \textit{id.} at 366-67.
\textsuperscript{153} E.g., Miguel A. Mendez, California’s New Law On Character Evidence: Evidence Code Section 352 and the Impact of Recent Psychological Studies, 31 U.C.L.A. L. REV. 1003, 1052-53 (1984); Robert G. Lawson, Credibility and Character: A Different Look at an Interminable Prob-
Despite the arguments raised in the legal literature, Mischel’s critiques have not led to the abandonment of the concept of personality. Much recent research reaches contrary conclusions to Mischel’s regarding the predictive value of the personality concept. Still other research concludes that individuals differ in the degree to which their behaviors display consistency as well as in what characteristics are likely to be consistent; thus “each of us possesses particular [personality] traits that are resistant to situational influence and others that can be readily modified.”

Mischel’s work has led to a rethinking of how to define basic personality characteristics. Such an approach may focus on “coherence,” that is, on identifying the logical patterns or regularities underlying behavior rather than on the sameness of the behavior itself. For example, excessive timidity in one context and extreme aggression in another may demonstrate that an individual has “a strong conflict over aggression.” Even more importantly, some theorists have recognized that traits need to be defined more precisely in terms of situations. Thus, instead of describing someone as “conscientious,” suggesting that the person is conscientious about all things, a person might be described as a “conscien-

---

lem, 50 Notre Dame Law. 758, 783 (1975) (discussing theories contrary to the notion of “stability”).
154. See infra note 162 and accompanying text.
155. See Millon, supra note 152, at 21-22.
156. See id. at 22-23.
157. Id. at 23.
158. Id. at 21-22. One modern moralist has defined a “moral character trait” as including both dispositions to do certain acts and to feel appropriate moral motivations. This motivational view of character is similar to the psychologist’s concept of “coherence”:

[If we use the broader motivational view of character as our baseline, a far smaller category of actions will be described as morally “out of character” even if the person had never acted like that before. If we have a reasonably complete picture of a person’s moral character, most of his moral and immoral future actions will reflect some enduring moral character trait—even if what we learn is new information about how that trait motivates behavior under new and unusual circumstances.

159. Millon, supra note 152, at 22 (quoting P.L. Wachtel, Psychodynamics, Behavior Theory and the Implacable Experimenter: An Inquiry into the Consistency of Personality, 82 J. Abnormal Psychol. 324 (1973)).
160. See Maddi, supra note 142, at 578 (“There is ample precedent among personality theorists . . . for a conceptualization of peripheral characteristics as potentialities that await situational cues for arousal.”).
tious student."\textsuperscript{161}

Indeed, Mischel himself has come around to the "interactionist" position, which reformulates the study of behavior as seeking an answer to this question: "Which personality and situational factors interact to produce consistent behaviors and which to produce variable behaviors?"\textsuperscript{162} A battered woman might, therefore, be especially fearful of violent assaults by her husband but not necessarily unusually fearful of all people, or even of all men.\textsuperscript{163} The need to link general traits to the specific situation before the jury is therefore critical in assessing the value of psychological character testimony.

**B. Objective Personality Tests**

1. *Creation and Measurement.*—"Objective" personality tests are those that "consist of true-false or multiple choice items that can be scored objectively."\textsuperscript{164} Different strategies may be used to design different tests, although strategies are often combined.\textsuperscript{165} One common strategy, the "rational-theoretical" strategy, deduces ap-

\textsuperscript{161} See Walter Mischel, Alternatives in the Pursuit of the Predictability and Consistency of Persons: Stable Data That Yield Unstable Interpretations, 51 J. Personality 578, 592-93 (1983) (discussing research regarding student conscientiousness). Accord, Steven J. Sherman & Russell H. Fazio, Parallels Between Attitudes and Traits as Predictors of Behavior, 51 J. Personality 308, 309 (1983) ("It may be, of course, that Tom is aggressive only at work or only toward men or that Alan is empathetic only during spring, but this is a matter of situational specificity of the expression of the trait.").

\textsuperscript{162} MILLON, supra note 152, at 22 (emphasis added). See generally Sherman & Fazio, supra note 161 (surveying research regarding the interrelationship between personality traits, attitudes, and situations).

\textsuperscript{163} Kinports, supra note 47, at 416 (noting the extreme fear that a battered woman may have of her husband).

\textsuperscript{164} Aiken, Assessment, supra note 143, at 192. The objective personality tests described in this Article involve the "nomothetic" approach to personality study: the search for general laws of behavior and personality by relying on group norms and averages. Id. at 423. An alternative method is the "idiographic approach," in which each individual "is considered to be a lawful, integrated system to be studied in his or her own right." Id. Although the idiographic concept has a long history and there has been some promising recent work in the area, see id. at 423-24, the approach has been challenged as "unscientific." Hergenhahn, supra note 144, at 197. But see id. at 245 (noting B.F. Skinner's support for idiographic research); infra text accompanying notes 385-405, 546-562 (arguing that there is a role for expert "artistry," as well as science, at a criminal trial). On the other hand, it has also been suggested that single subject research, as opposed to group research, can aid both in understanding unique individuals and in better formulating laws of human behavior. Marshall Edelson, Hypothesis and Evidence in Psychoanalysis 63-69 (1984). In any event, "idiographic" research is rarely the explicit subject of the relevant case law and is often given inadequate attention in the standard texts on personality assessment and theory. See, e.g., Aiken, Assessment, supra note 143; Hergenhahn, supra note 144. These are oversights that this Article will address. See infra text accompanying notes 385-404, 546-562.

\textsuperscript{165} See Aiken, Assessment, supra note 143, at 194.
propriate questions from logic and a particular psychological theory. An alternative common strategy is more empirical: the "criterion-group" method. This method begins with a sample with known characteristics, such as a group of diagnosed schizophrenics. An item pool is then administered to individuals in the known sample and to a control group—usually a normal population. The items that distinguish the known sample from the normal group are then placed in a scale. Those items are then used on another similar sample to determine whether the scale continues to distinguish between the two groups.

Two concerns guide judgments about the value of a particular personality test: reliability and validity. "Reliability" refers to the consistency of the testing device, that is, its ability to arrive at the same conclusion from tests given at different times and in different situations, assuming that there is no underlying change in the trait being measured. There are many different ways of measuring reliability, but two notable ones are retesting the same group, and having different examiners score the same test. A "correlation coefficient," a statistical measure of reliability, is then assigned.

"Validity" reflects the extent to which the test measures what it was created to measure. Here, too, there are different ways to make this determination. "Criterion-related validity" measures test scores against an external criterion. If that external criterion is measured at the same time as the test is given, that is called "concurrent validity." For example, does the test classify mental patients in the same category as do psychiatrists using other techniques? If the external criterion is measured after the test is

166. Id. at 192-93.
167. Id. at 194.
169. AIKEN, ASSESSMENT, supra note 143, at 67.
170. Id. at 73.
171. Id. at 74-77.
173. AIKEN, ASSESSMENT, supra note 143, at 81.
174. Id. at 82, 466.
175. Id.
176. See id.
given, that is called "predictive validity." 177 For example, does a measure of violence correspond to later observations of violent behavior?

Understanding predictive validity requires an attention to "base rates," the rate of the observed characteristic in the overall population. 178 For instance, it may be that sixty percent of the American adult population does not commit violent acts. A test resulting in accurate predictions of nonviolence sixty percent of the time thus tells us nothing that we did not already know. 179 Therefore, one must also consider "incremental validity," that is, whether the test is a more accurate measurement over an alternative diagnostic tool. 180

Also of importance is "construct validity": whether the test measures a particular psychological construct, such as anxiety or achievement motivation. 181 Expert judgments about whether a test has such validity, analysis of the test's internal consistency, and questioning of examinees after tests to reveal mental processes in deciding to answer questions in a certain way, are all considered in gauging construct validity. 182 To have construct validity, a test should have high correlations with measures of the construct by other means and low correlations with measures of different constructs. 183

Several further cautions are necessary. First, standard procedures should be followed in administering tests to minimize the effects of the physical surroundings and the examiner's behavior on the person examined. 184 Second, "generalization" must be done cautiously. 185 Thus, if a test's validity is based on a sample of all white, middle-class males, results from administering the test to members of different racial and age groups may not lead to the same

177. Id. at 82, 473. See also P.E. Vernon, The Concept of Validity in Personality Study, in PERSONALITY ASSESSMENT 407, 408 (B. Semeonoff ed., 1966) (mentioning prediction of "legal offenses"); AIKEN, ASSESSMENT, supra note 143, at 82-83 (offering additional examples of predictive validity).

178. AIKEN, ASSESSMENT, supra note 143, at 84, 465.

179. These figures are a variation on an example offered in ZISKIN & FAUST, supra note 172, at 548.

180. See AIKEN, ASSESSMENT, supra note 143, at 84; ZISKIN & FAUST, supra note 172, at 549.

181. AIKEN, ASSESSMENT, supra note 143, at 85.

182. Id.

183. Id. at 85-86.

184. See id. at 93-102 (outlining procedures and their importance); ZISKIN & FAUST, supra note 172, at 543-44 (stressing the importance of uniformity in administering tests).

185. See infra notes 186-187, 408-411.
conclusions. Moreover, if a test measures one particular trait, explanations extending that trait to other characteristics and circumstances are suspect. Indeed, concerns about this problem have led some to advocate the creation of a wide battery of tests, each of which is highly specific to a characteristic important in a courtroom setting; early efforts on this score are promising.

2. A Brief Comment on the MMPI.—It is worth briefly discussing the Minnesota Multiphasic Personality Inventory (MMPI) as an example of an objective test, because it is perhaps the most widely used psychological instrument, especially in court-related settings. Over 12,000 research articles and books have been written about the MMPI since it was developed over 50 years ago.

The MMPI consists of 566 statements to be classified as true or mostly true, or false or not usually true. The MMPI includes four “validity” scales, the primary purpose of which is to detect invalidity in a particular administration of the exam because the subject may have answered in such a way as to be perceived in a more or less favorable light, or been confused or overly critical and defensive in

186. See Ziskin & Faust, supra note 172, at 550-51 (discussing the extent to which test results can be generalized); Lynne B. Rosewater, The Clinical and Courtroom Application of Battered Women’s Personality Assessments, in DOMESTIC VIOLENCE ON TRIAL: PSYCHOLOGICAL AND LEGAL DIMENSIONS OF FAMILY VIOLENCE 87 (Daniel J. Sonkin ed., 1987) (noting that because control group used in creating MMPI, an objective personality assessment test, contained no blacks, most experts agree that accurate interpretation of the test had to take into account the race of the test-taker). However, “it is often reasonable to assume that if the changes in persons or circumstances are very minor, there is a good chance to expect generalization.” Ziskin & Faust, supra note 172, at 550.

187. See Ziskin & Faust, supra note 172, at 550 (“A test that may help measure skills essential to a career in accounting (e.g., mathematical abilities) probably will not help in the assessment of marital accord.”).

188. See Richard I. Lanyon, Psychological Assessment Procedures in Court-Related Settings, 17 PROFESSIONAL PSYCHOLOGY: RESEARCH AND PRACTICE 260-68 (1986) (discussing movement toward use of issue-specific assessment tools); Aiken, Assessment, supra note 143, at 437 (noting special tests for competency, types and strengths of sexually anomalous behavior, and a child’s perception of her parents). In addition to increased research on issue-specific tests, there has been a growing focus on devising tests to assess the dynamic, two-way interaction between persons and environment discussed supra text accompanying notes 142-163. See Lewis Aiken, PSYCHOLOGICAL TESTING AND ASSESSMENT 434-36 (7th ed. 1991) [hereinafter Aiken, Testing] (summarizing many of the new tests focusing on this interaction).

189. Aiken, Testing, supra note 188, at 433 (stating that the MMPI is perhaps the “most frequently administered of all psychometric instruments in legal settings”); Daniel Shuman, Psychiatric and Psychological Evidence 54 (1986) (stating that the MMPI is the “premier diagnostic and screening device in clinical psychology”).

190. Marc J. Ackerman & Andrew W. Kane, How to Examine Psychological Experts in Divorce and Other Civil Actions 196 (1990).

191. Aiken, Assessment, supra note 143, at 231.
responding. The remaining scales are clinical scales from which traditional psychiatric diagnoses are made and research or supplementary scales designed to measure specific personality characteristics or measure specific problem areas. However, a psychiatric diagnosis or personality analysis is based not on a single scale score but on a combination of scores. The scores are interpreted by comparing the results to norms and to the many published research findings and by using professional judgment to determine possible explanations for tendencies shown in the results, based on the subject’s background. The reliability coefficients vary for different scales, with median coefficients ranging from .7 to .9.

One critical focus of inquiry in using the MMPI in court is to examine any available empirical evidence concerning its use for the particular purpose served in bringing the psychologist to court. For example, specific research has been done examining the MMPI results of battered women. Unfortunately, such highly specific empirical research is often lacking.

The MMPI has recently been revised, being replaced by the MMPI-2. The test was revised because some of the items were out of date, some were objectionable as sexist or otherwise offensive, some were not scored on useful scales, and there was a need to expand the item pool to address new problem areas and to develop

192. Id. at 234-35.
193. ACKERMAN & KANE, supra note 190, at 154. See also AIKEN, ASSESSMENT, supra note 143, at 233-34 (describing the clinical scales, including the “depression,” “psychopathic deviate,” “paranoia,” and “hypomania” scales).
194. AIKEN, ASSESSMENT, supra note 143, at 236.
195. ACKERMAN & KANE, supra note 190, at 155.
196. AIKEN, ASSESSMENT, supra note 143, at 239. The reliability coefficients vary based on the approach used in computing the coefficient. These relatively high figures were computed by the method of “internal consistency coefficients.” Id. A wider range, from .58 to .92 has been noted for the clinical scales on the revised MMPI when a different method, test-retest, is used to compute reliability. AIKEN, TESTING, supra note 188, at 378.
197. See Rosewater, supra note 186, at 89-91. Professor Rosewater found that the highest scale elevation for battered women was on the scale measuring anger. Id. at 89. Subscales showed, however, that battered women usually direct their anger inward instead of outward. Id. They also show elevation on the scales measuring confusion and paranoia or fearfulness. Id. But intactness and ego strength scales were low, a finding consistent with the theory that battered women suffer from learned helplessness. Id. at 90.
198. E.g., AIKEN, TESTING, supra note 188, at 377 (“To provide a more representative sample of the U.S. adult population than the original MMPI, MMPI-2 was standardized on 2,600 U.S. residents aged 18 to 90 ... selected according to statistics provided by the 1980 census on geographic distribution, ethnic and racial composition, age and educational levels, and marital status.”); ACKERMAN & KANE, supra note 190, at 196.
new norms.\textsuperscript{199}

\section*{C. Projective Tests}

Projective techniques consist of "relatively unstructured stimuli that examinees are asked to describe, tell a story about, complete, or respond to in some other manner."\textsuperscript{200} Two examples are the Rorschach Ink Blot Test and the Thematic Apperception Test.\textsuperscript{201} In the Rorschach Test, subjects are shown inkblots and asked to describe what the inkblots suggest to them.\textsuperscript{202} There are various methods of administration and scoring.\textsuperscript{203} In the Thematic Apperception Test, the subject is shown a series of cards, most with human figures on them, and is asked to make up a story about the picture.\textsuperscript{204}

Supporters of projective tests maintain that the structure imposed on the inkblot or picture reflects fundamental facets of personality.\textsuperscript{205} But even these supporters admit that the tests are highly fakable, are very susceptible to the conditions under which they are administered, are scored and interpreted differently by different examiners, and have inadequate or unrepresentative norms or no norms at all.\textsuperscript{206} Indeed, these features limit the value of these tests as a useful way of communicating information in a courtroom, and a majority of clinical psychology faculty in universities view the tests negatively.\textsuperscript{207}

\section*{D. Observations and Interviews}

In "observation," the expert observer simply takes note of and perhaps records the behaviors displayed by a subject under

\begin{footnotesize}
199. Aiken, Testing, supra note 188, at 377.
200. Aiken, Assessment, supra note 143, at 305.
202. Id.
203. Id. at 167.
204. Id.
205. See Aiken, Assessment, supra note 143, at 306 ("The more unstructured the task, the more likely the responses are to reveal important facets of personality.").
206. Id. at 307. But see Ackerman & Kane, supra note 190, at 234-37, 241-45 (arguing that Rorschach has greater reliability and validity than is commonly suggested and that both the Rorschach and the Thematic Apperception Test are useful as interviewing aids to be combined with other techniques); E. Jerry Phares, Clinical Psychology: Concepts, Methods, and Profession 224 (4th ed. 1992) (suggesting that the Rorschach may be a useful instrument when an adequate theoretical or empirical basis exists for making specific predictions).
\end{footnotesize}
Observations can be uncontrolled but systematic, as where specific, critical behaviors are identified for the observer to record. In controlled observations, on the other hand, the subjects are placed in prearranged or contrived situations with the goal of observing how they behave in those circumstances. The accuracy of observation can be improved by requiring immediate recording of what is observed, multiple observers, videotaping, and careful training of observers regarding what to watch for and how to minimize the impact of the observer's own biases, expectancies, personality, attitudes, or desires.

With interviews, instead of merely observing, the psychologist asks questions and records answers. Interviews can be unstructured, with the interviewer free to ask what seems appropriate, or structured, with a series of specific, standard questions to be asked. Answers from the latter type of interview can sometimes be quantified. Interviews may cover such topics as physical appearance; behavior; family, medical, educational, marriage, and employment history; interests and attitudes; and current problems. Carefully planned, structured, and specific interviews conducted by specially trained interviewers are more likely to be valid. Validity may also be improved by using multiple evaluators, which can be facilitated by the use of videotapes. In addition, some experts recommend lengthy, multiple interviews performed by different interviewers in different settings.

Forensic psychologists also seek out sources other than the defendant concerning the offense, previous antisocial behavior, general historical information, and the relevant psychiatric information. This helps to verify the defendant's "story" and

---

208. *Aiken, Assessment*, supra note 143, at 128.
209. *Id.* at 128-29.
210. *Id.* at 129.
211. *Id.* at 132-33.
212. See *id.* at 140-41 (describing when to use structured and unstructured interviews).
213. *Id.* at 141.
214. *Id.* at 141-42.
215. *Id.* at 150.
216. See infra text accompanying notes 384-404.
provides the therapist with information that the defendant may not have. This may require interviewing many people who play an important role in the defendant's life.

E. Diagnoses and DSM-III-R

The revised third edition of the Diagnostic and Statistical Manual of the American Psychiatric Association (DSM-III-R) provides a means of classifying psychological and psychiatric disorders. The DSM-III-R may improve the reliability of psychological diagnosis by making diagnostic criteria more concrete, behavioral, and observable than had been true in the past. Videotaping of interviews to train clinicians to use the system uniformly has also helped to increase reliability. Diagnoses using the DSM-III-R are "probabilistic" in that patients with a given disorder can have quite different signs and symptoms because, for example, a patient can qualify for the same diagnosis by having any three symptoms from Group A, any two from Group B, and none from Group C. Diagnoses are accepted or rejected after combining the observed symptoms according to a prescribed formula. The diagnosis is applied if the critical number of symptoms in an acceptable combination is reached.

DSM-III-R diagnoses may be of greater importance in psychiatric testimony requiring that the defendant suffer from a mental disease or defect, as is true with the insanity defense and diminished capacity, because the diagnosis supports the argument that the de-

219. Id. at 508-09.
221. HOOD & JOHNSON, supra note 168, at 179.
222. SHUMAN, supra note 189, at 30. But see Nancy Cantor, "Everyday" Versus Normative Models of Clinical and Social Judgment, in INTEGRATIONS OF CLINICAL AND SOCIAL PSYCHOLOGY 27 (Gifford Weary & Herbert L. Mirels eds., 1982) (critiquing the traditional method of psychiatric diagnosis and suggesting an alternative approach that more closely mirrors the way in which real clinicians behave).
223. SHUMAN, supra note 189, at 30.
224. PAUL R. MCHugh & PHILLIP R. SLAVNEY, THE PERSPECTIVES OF PSYCHIATRY 30 (1986). Although doctors McHugh and Slavney's description was under DSM-III, the predecessor to DSM-III-R, this probabilistic feature has not changed in the revised manual. See, e.g., AMERICAN PSYCHIATRIC ASSOCIATION, DESK REFERENCE TO THE DIAGNOSTIC CRITERIA FROM DSM-III-R 49 (1987) (describing "autistic disorder" as requiring the presence of at least eight of sixteen listed items, "to include at least two items from Group A, one from Group B, and one from Group C").
226. Id.
Fendant indeed suffers from a mental disease.\textsuperscript{227} For psychological character evidence, however, a DSM-III-R diagnosis may be helpful\textsuperscript{228} but is not essential, because it is not necessary that the defendant suffer from an illness or disorder at all.\textsuperscript{229}

\textbf{F. Profiles, Syndromes, and Typologies}

A psychologist may also rely on empirical studies regarding common traits of members of certain groups. The characteristics of group members are often described as "syndromes," "profiles," or "typologies," terms among which, for purposes of this Article, there are no real distinctions.\textsuperscript{230} These profiles may differ in the extent to which there is supporting empirical research and agreement regarding their validity among the scientific community. For example, many psychologists agree that there is significant support for the battered woman syndrome.\textsuperscript{231} But there may be less agreement on whether there are rapist typologies.\textsuperscript{232} Beyond this, some psychologists have suggested that reasonable working "hypotheses" can be formulated regarding the "type" of person who would commit a particular crime, based on specific personality characteristics required to commit the crime.\textsuperscript{233} For example, sustained fraud generally requires high intelligence.\textsuperscript{234} Regardless of the source of the profile, it must be compared to the results of the clinician's other investigations regarding the patient to determine whether the indi-

\begin{itemize}
\item \textsuperscript{227} See infra text accompanying notes 575-583.
\item \textsuperscript{228} McCord, Profiles, supra note 119, at 99 (noting that the availability of "clear diagnostic criteria" bolsters the "plausibility" of proffered psychological evidence).
\item \textsuperscript{229} See infra text accompanying notes 403-404, 575-583.
\item \textsuperscript{230} See, e.g., McCord, Profiles, supra note 119, at 24 n.14 (defining syndromes and profiles); Hans Toch, Violent Men: An Inquiry into the Psychology of Violence 175-77, 214 (1969) (suggesting a "typology" of the violence-prone person but questioning its validity).
\item \textsuperscript{232} Compare Samuel Roll & William E. Foote, The Inconsistent Personality Defence, in Psychology and Law: Topics From an International Conference 125, 128-30 (David J. Müller et al. eds., 1984) (favoring use of "rapist typologies" to support a defendant's argument that his personality was "inconsistent" with rape) with Lanyon, supra note 188, at 265 (noting that some reviews have shown that "sex offenders in general do not have backgrounds and personality characteristics that make them reliably distinguishable").
\item \textsuperscript{233} See Roll & Foote, supra note 232, at 126-27 ("[C]ertain crimes (e.g., sustained fraud) directly point to some aspects of the person (capacity to plan, moderately high to high intelligence)."). Roll and Foote are thus among those theorists who view intelligence as an aspect of personality. See supra text accompanying notes 145-148.
\item \textsuperscript{234} See Roll & Foote, supra note 232, at 126-27.
\end{itemize}
individual fits the profile. 235

G. Case Studies and Life Stories

These interviews, observations, and other investigations are combined with the personality tests and reviews of the empirical literature to produce a "case study." 236 The case study seeks to create an overall picture of the individual, relying upon multiple sources of information and the consistencies or inconsistencies between those sources. 237 Case studies inevitably include some subjective factors, but the extent to which these factors result in biases can be reduced by carefully training clinicians to avoid potential sources of bias. 238 Additionally, careful recording of all "objective" data forming the basis for the clinician's opinion is necessary. 239 Ways to improve the quality of case studies and their usefulness to a jury are discussed in greater detail below. 240

The point of the case study, of course, is to tell the individual patient's life story. 241 To be useful to the jury, it must be built upon as much objective data and empirical research as possible. But it also must be a story that is unquestionably based to a large degree on the experience and judgment of the clinician, for much of what psychologists rely on in their daily practice has not yet been adequately empirically validated. 242 Psychological opinions should,
therefore, be viewed as containing both scientific and "artistic" components.\textsuperscript{243} It is this dual nature of psychological evidence that makes for particularly thorny questions of admissibility in criminal trials—questions to which this Article now turns.

III. Evidentiary Analysis

The analysis thus far has focused on two broad themes: (1) that there are both moral and social psychological reasons favoring the admissibility of psychological character evidence to promote individualized justice in criminal prosecutions; and (2) that the scientific and artistic bases for such evidence, and an understanding of precisely what expert character witnesses do, are essential for making admissibility decisions in particular instances. These themes will guide much of the evidentiary analysis of pragmatic relevancy to be addressed below. That analysis must begin by examining Federal Rule of Evidence 404 and the related special rules governing the admissibility of character evidence, rules that were crafted to standardize practices and address concerns about the probative worth and prejudicial impact of character evidence, that is, about its pragmatic relevancy.

A. Federal Rule of Evidence 404

The law of evidence demonstrates a fundamental distrust of much character evidence.\textsuperscript{244} That distrust is embodied in Federal Rule of Evidence 404, which declares in part that "[e]vidence of a person's character or a trait of character is not admissible for the purpose of proving action in conformity therewith on a particular occasion."\textsuperscript{245} However, the rule creates certain exceptions, the most important of which has been called the "mercy rule," which permits an accused in a criminal case to offer evidence of a "pertinent" trait of character for exculpatory or mitigation purposes; the

\begin{itemize}
\item that valid studies using other theories regarding the problems of battered women are available).
\item 243. See, e.g., LESTON HAVENS, APPROACHES TO THE MIND: MOVEMENT OF THE PSYCHIATRIC SCHOOLS FROM SECTS TOWARD SCIENCE 322 (1987) ("[P]sychartry is a practical art with scientific aspirations . . . "). \textit{But see generally} EDSELON, supra note 164 (rejecting claims that psychoanalysis is not a science and suggesting ways to improve its scientific standing).
\item 244. See, e.g., MCCORMICK ET AL., supra note 16, § 186 (stating that while character evidence "almost always has some probative value, in many situations, the probative value is slight and the potential for prejudice large"); ZUCKERMAN, supra note 95, at 234-37, 241-46.
\item 245. FED. R. EVID. 404(a).
\end{itemize}
prosecution would then be permitted to offer character evidence in rebuttal. If there is reason to distrust character evidence, however, why should it be treated differently when first offered by the accused instead of the prosecution? This Article has suggested one plausible justification: it promotes, in furtherance of moral and social psychological concerns, the reality and the perception of individualized justice, even when there are reasons to be concerned about the value and impact of the evidence.

The more difficult question is when would prosecution-proffered evidence be considered evidence "of a person's character" within the meaning of Rule 404. It is clear that Rule 404's bar is directed, at most, against the circumstantial use of character evidence—the argument that because a person has a particular trait, that person was more (or less) likely than someone without that trait to act or think in a particular way. That Rule 404 does not seek to bar direct use of character evidence—that is, proving character when it is itself an element of a crime—is apparent from Rule 405, which permits proof of character when character is indeed an "es-

246. See Fed. R. Evid. 404(a)(1); Weisbenner, supra note 18, § 404.6 (noting that admissibility of good character evidence for the defense embodies a "mercy principle"). Rule 404 also makes exceptions for certain evidence as to the character of the victim, Fed. R. Evid. 404(a)(2), but this Article addresses only evidence of the defendant's character. Rule 404 makes a further exception for evidence of the character of a witness as to truthfulness or untruthfulness as permitted by Rules 607-09. Fed. R. Evid. 404(a)(3), 607-09. While a defendant may be a witness, the special problems of a witness's character for truthfulness or untruthfulness have been addressed effectively elsewhere and need not be repeated here. See, e.g., McCord, Profiles, supra note 119, at 44-48, 76-79; Mosteller, supra note 126, at 105-09. Much of the discussion here, however, regarding how to measure the value of any expert character witness would, of course, also apply to analyzing psychological character testimony as to a witness's truthfulness.

247. See supra text accompanying notes 83-118 and infra text accompanying notes 330-344. Cf. Zuckerman, supra note 95, at 22-23, 235-36 (suggesting that the "mercy principle" may result in acquittal based upon a jury's assessment of a defendant's moral worthiness). Prosecution-proffered character evidence may sometimes help to individualize justice as well, see supra text accompanying note 117 and infra note 664; but for reasons discussed below, defense claims to individualized justice are entitled to greater weight. See infra text accompanying notes 336-345. Differential treatment of prosecution and defense thus makes sense. Nevertheless, because prosecution evidence may sometimes individualize justice and will not necessarily prejudice the jury, it is debatable whether prosecution first-use of expert character evidence should flatly be barred in all cases. See infra note 664. Another justification for differential treatment is that a defendant's use of evidence of good character is thought to be less prejudicial to the jury and may have fewer social costs (creating a reasonable doubt as to guilt) than prosecution-offered evidence of a defendant's bad character. See McCormick et al., supra note 16, § 191.

248. See Imwinkelried, supra note 19, at 575-80 (discussing two potential circumstantial uses for character evidence).
sentential element of a charge, claim, or defense." 249 Such proof is direct, not circumstantial, and thus not barred by Rule 404.250

The currently prevailing view, however, is that Federal Rule of Evidence 404 does not reach as far in barring circumstantial uses of character evidence as the rule might logically permit. Instead, the rule is viewed as barring prosecutorial use of character evidence only when it would be used circumstantially to prove a defendant’s identity or behavior, but not when it would be used to prove a defendant’s mental state.251 Such an interpretation is indeed consistent with the plain meaning of the rule, which bars the use of character to prove "'action in conformity therewith.'"252 Yet such a narrow conception makes little sense, for both types of circumstantial use, to prove behavior and to prove mental state, raise similar dangers.253

Circumstantial use of psychological character evidence to prove behavior involves the following inferential chain: (1) the psychologist offers her opinion, and perhaps supporting data and theories, that the defendant has a certain personality trait; (2) that opinion supports the inference that the defendant indeed has that trait; and (3) someone with that trait is more likely than someone without it to commit the evil deed.254

Proving mental state circumstantially by using character testimony involves substantially the same inferential chain. The only difference is that the third step involves an inference that the trait makes it more likely than otherwise that a defendant had a certain state of mind, not that a defendant performed a certain deed.255 As with direct use of character evidence, both circumstantial uses raise concerns that juries will convict defendants solely for who they are, not for what they have done,256 and that psychologists’ opinions

249. Fed. R. Evid. 405(b). For example, under a subjective theory of entrapment, defendants must prove that the government induced them to commit a crime that they were not otherwise predisposed to commit. Wayne R. LaFave & Jerold H. Israel, Criminal Procedure 280 (2d ed. 1992). Arguably, entrapment requires proof of a character trait—the absence of a "predisposition" to commit the crime—as an element of the offense. See, e.g., United States v. Hill, 655 F.2d 512, 516 (3d Cir. 1981).

250. See United States v. Rippy, 606 F.2d 1150, 1155 n.34 (D.C. Cir. 1979).

251. Id.

252. Id. at 578 (quoting Fed. R. Evid. 404(b)).

253. Id. at 582-83.

254. Cf. id. at 576-579 (tracing similar inferential chains to those discussed here, but using specific acts, instead of expert opinion, as the type of character evidence offered).

255. See id. at 579.

256. Carlson et al., supra note 21, at 449 (noting that a prosecutor's argument that Devitt, an accused rapist, was an "immoral, law-breaking person" creates a "grave risk
that defendants have the noted trait are unreliable. Both circum-
stantial uses raise one additional concern: a jury may give too much
weight to the existence of a personality trait as a predictor of a de-
fendant's actions or thoughts in any particular instance. These
similar policy concerns suggest that both types of evidence should
be subject to a Rule 404-type bar when offered by the prosecu-
tion. Most courts, however, do not automatically bar character
evidence to prove mental state.

Several consequences flow from these observations. First, Rule
404 alone will not, in the view of most courts, bar admission of char-
acter evidence when used to prove mental state, even though the
character evidence raises many of the policy concerns toward which
Rule 404 is directed. Consequently, defense counsel must be
prepared to attack prosecution use of such evidence on other
grounds, the most promising of which relate to pragmatic rele-
vaney.

Second, some courts seem to view the mercy rule as effectively
creating an automatic defense right to use expert character evi-
dence. The mercy rule does permit defense use of character evi-
dence under Rule 404, and Rule 405 permits proof of character by
reputation or opinion. Moreover, the comments to Rule 405 specif-
ically recognize that the drafters contemplated the admissibility of

of a decision on an improper basis—that is, the jury's belief that Devitt is such an antiso-
cial character that he belongs behind bars even if he happens to be innocent of the rape with
which he is now charged”) (emphasis added).

257. See infra text accompanying notes 293-310.
258. See Imwinkelried, supra note 19, at 580-84. This last concern is not present with
direct use of character evidence, because then the sole concern is proving that the de-
fendant exhibits a specific character trait. Once the trait is proven, the inquiry is at an
end. An analysis of whether that trait made it more or less likely that a defendant would
behave in a particular way would be irrelevant.

259. On the other hand, despite these similar policy concerns, there is an argument
that with respect to psychological character evidence, a flat bar is inappropriate. See infra
note 664 and accompanying text.
260. Imwinkelried, supra note 19, at 579.
261. Id. at 582-83.
262. See infra notes 347-350 and accompanying text.
that the lower court erred in excluding expert psychiatric testimony to establish that the
defendant had the character trait of impulsiveness from which the jury could have con-
cluded that he did not premeditate the murder); United States v. Hill, 655 F.2d 512,
516-17 (3d Cir. 1981) (holding that expert testimony to prove that the defendant was
not predisposed to commit the crime and therefore was entrapped is admissible under
Rules 702 and 703); State v. Rives, 407 So. 2d 1195, 1198-99 (La. 1981) (expressing
concern over the trial court's refusal to admit character evidence to establish good repu-
tation as to pertinent traits).
expert character opinion. Some courts, however, have seen these two rules, or their state law analogues, as creating a presumptive right for the defense to submit expert character evidence, with little real inquiry into the value or jury impact of that evidence.

There may indeed be strong reasons to weigh the pragmatic relevancy balancing in favor of defense-offered character evidence; there is little justification, however, for ignoring an inquiry into the accuracy and impact of that evidence, and nothing in any traditional evidentiary doctrine justifies such an approach. This is also true for the prosecution, which must be able to articulate clear arguments based on the probative value and potential prejudicial impact of expert character testimony.

B. The Frye Rule

The classic test for the admissibility of scientific evidence was articulated in Frye v. United States:

Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.

Over the years, there has been much dispute over the validity of the Frye test. Many jurisdictions nevertheless continue to apply

---

264. If character is defined as the kind of person one is, then account must be taken of varying ways of arriving at the estimate. These may range from the opinion of the employer who has found the man honest to the opinion of the psychiatrist based upon examination and testing.

FED. R. EVID. 405 advisory committee's note (emphasis added).

265. See Christensen, 628 P.2d at 582; Rives, 407 So. 2d at 1198. Cf. United States v. Staggs, 553 F.2d 1073, 1075 (7th Cir. 1977) (suggesting that FED. R. EVID. 404 alone creates a right to admit pertinent defense-proffered psychological character evidence); State v. Treadaway, 568 P.2d 1061, 1066 (Ariz. 1977) (suggesting FED. R. EVID. 405 creates a presumptive right to the admission of defense character evidence, whether or not the witness has any special psychological expertise).

266. 295 F. 1013 (D.C. Cir. 1923).

267. Id. at 1014.

268. E.g., Paul C. Giannelli, Frye v. United States: Background Paper Prepared for the National Conference of Lawyers and Science, 99 F.R.D. 189, 191-92 (1983) ("Notwithstanding its widespread judicial adoption, the general acceptance test remains controversial. A number of courts have rejected it and commentators have labeled it 'archaic,' 'a sport,' 'infamous,' and 'antiquated on the day of its pronouncement.'" (footnotes omitted)).
Frye as a potential obstacle to admissibility.269 Even in jurisdictions recognizing Frye, however, the test should properly be understood to apply only to evidence that is both “novel” and “scientific,” two limitations often forgotten or inadequately addressed.270 Two of the leading decisions on the question of whether Frye applies to psychological character evidence illustrate the strengths and weaknesses of the courts’ handling of Frye in this area.

In the first of these cases, State v. Cavallo,271 the trial judge in a rape case barred the testimony of Cavallo’s psychiatrist, who would have testified that Cavallo did not have the psychological traits of a rapist.272 The appellate court affirmed, concluding that the state’s Frye equivalent273 was applicable and had not been met.274 The court applied the Frye test for one primary reason: the court feared that jurors would be so impressed by the testimony that they would give it undue weight, despite its lack of general acceptance among other experts in the field.275 This distrust of the jury was fundamental to the decision. Merely labeling testimony “expert,” the court believed, would deeply and unduly impress the jury.276 Moreover, according to the court, in so “subjective a field as psychiatry,” a


271. 443 A.2d 1020 (N.J. 1982).

272. Id. at 1022. The psychiatrist also would have testified about Cavallo’s individual qualities, as opposed to whether his qualities matched those of another group. Specifically, he would have testified that Cavallo was a “well-meaning individual, he would not wilfully do a wrong, he recognizes the force and violence of rape are wrongful acts, [and] he is a non-violent, non-aggressive person.” Id. at 1021-22. Nevertheless, the focus of concern in the court’s opinion was on the overt group comparison, namely that Cavallo lacked the characteristics typical of the group labeled “rapists.” See id. at 1027.


274. See Cavallo, 443 A.2d at 1022, 1026. The court concluded that the proffered testimony was “expert character testimony,” permissible under New Jersey’s character evidence rules, which are similar to the Federal Rules of Evidence. Id. at 1023. Nevertheless, the court emphasized that a proper foundation first must be laid to demonstrate that the testimony was proper expert evidence; thus, the court essentially applied the Frye test. Id. at 1023-24. The court suggested that New Jersey’s standard was in some undefined way different from the Frye standard but seemingly followed a standard Frye analysis. See id. at 1026.

275. See id. at 1024.

276. Id.
“battle of the experts” would surely confuse the jury and divert its attention from the central question of whether the defendant committed the rape. The court concluded that juries, unlike expert bodies such as parole boards, would be unable to judge and evaluate this dispute among experts fairly. Although juries were previously permitted to hear psychiatric testimony regarding mental illness or state of mind, the court saw the question of whether a psychiatrist could predict behavior as a novel use of psychiatric evidence that, given the potential adverse impact on the jury, required the safeguards provided by Frye.

Unlike the appellate court in Cavallo, the Supreme Court of California in People v. Stoll expressed greater faith in the capacities of jurors and in the usefulness of psychological expertise in proving or disproving criminal acts. Stoll involved four defendants tried for lewd and lascivious conduct involving seven young boys. One of the defendants, Grafton, sought to call a psychologist to testify that she: (1) showed no indication of sexual or other deviancy, (2) had a low indication of antisocial or aggressive behavior, and (3) was unlikely to have been involved in the events for which she was charged. The bases for the proffered testimony were two objective personality tests (one of which was the MMPI), interviews with Grafton, and the psychologist’s experience with similar individuals. The psychologist admitted, however, that the MMPI had only about a seventy percent chance of being right when administered to “normal” persons; he also acknowledged that the test was not valid when given to disturbed or psychotic individuals, and conceded, further,

277. Id.
278. See id. at 1029.
279. See id. at 1027-29. The court’s opinion is not quite as clear as the text suggests. First, in the portion of its opinion addressing whether the Frye standard was met, i.e., whether the technique involved had been generally accepted, the court emphasized the way in which Cavallo’s use of the evidence was different from other accepted uses of psychological evidence. Id. But that analysis is the same as saying that the application of the technique in the case before the court was “novel.” This would be pertinent only to whether a Frye question arises, not, as the court was doing, to whether Frye was met under Cavallo’s unique application.

Second, part of the court’s analysis concerning the jury’s competence to handle the proffered evidence, an inquiry again relevant only to whether Frye applies at all, appeared instead in the portion of the court’s opinion addressing general acceptance. Id. at 1025. Despite this confusion, the court seemed to recognize the need for, and engage in, the two distinct inquiries—novelty and jury impact—that govern whether the Frye limitation should apply.

280. 783 P.2d 698 (Cal. 1989).
281. Id. at 700.
282. Id. at 704-05.
that one of his own patients, who was an "admitted" child molester, had nevertheless tested within normal limits on the MMPI.\(^{283}\) He went on to explain, moreover, that the objective personality tests were only factors in his "subjective judgment," which was not "an absolute cut and dried scientific approach."\(^{284}\) The trial court excluded the psychologist's opinion under the California equivalent of \textit{Frye}.\(^{285}\)

The appellate court reversed, engaging, as did the court in \textit{Cavallo}, in an inquiry into the jury's ability to handle the testimony to be presented before it.\(^{286}\) Unlike \textit{Cavallo}, however, the \textit{Stoll} court saw no special feature of the proffered testimony that "effectively blindside[d] the jury."\(^{287}\) While the witness had faith in the objective test results and described the MMPI's "validity scales," these facts did not make the tests appear foolproof, for the acknowledged weaknesses in the test's statistical accuracy and its proven fallibility in the psychologist's own practice were to be brought to the jury's attention.\(^{288}\) Moreover, the psychologist's explanation that the objective tests were merely the springboard for a diagnosis relying on other, more "subjective" sources, such as patient interviews and case history, made it clear that the process was more a professional art than a science.\(^{289}\) The court rejected the argument that using objective personality tests to prove that a defendant did not commit an act, rather than to prove that she did not harbor a requisite mental state, was "revolutionary."\(^{290}\) The court saw "no legal or logical basis"\(^{291}\) for treating testimony regarding behavior any differently from psychological testimony regarding mental state—including standardized personality tests—that had been permitted in the past.

The contrasting approaches of the courts in \textit{Cavallo} and \textit{Stoll} raise four distinct problems, addressed below.

\begin{itemize}[<1.5ex]
    \item \textit{1. Is Expert Character Testimony "Scientific"?—}The general approach apparently followed by both the \textit{Cavallo} and \textit{Stoll} courts—evaluating the likely impact of the proffered evidence on the jury—is consistent with \textit{Frye}'s main purpose: "to screen out expert evidence
\end{itemize}

\(^{283}\) \textbf{Id. at 705.}\ns\(^{284}\) \textbf{Id.}\ns\(^{285}\) \textbf{Id. at 706. See \textit{People v. Kelly}, 549 P.2d 1240 (Cal. 1976).}\ns\(^{286}\) \textbf{Stoll, 783 P.2d at 712, 715.}\ns\(^{287}\) \textbf{Id. at 710.}\ns\(^{288}\) \textbf{Id. at 712.}\ns\(^{289}\) \textit{See id.}\ns\(^{290}\) \textbf{Id.}\ns\(^{291}\) \textbf{Id.}
that the jury cannot properly evaluate." This inability usually results from a jury's tendency to be "overawed" by certain types of experts. Under Frye, therefore, expert evidence that "overawe" will be admitted only if there is an independent guarantee of the evidence's trustworthiness. Frye provides that guarantee by ensuring that the only expert testimony admitted results from techniques whose accuracy has been reviewed and accepted by a "minimal reserve [of] unbiased experts"—that is, by those experts constituting a substantial portion of the relevant scientific community. Any expert evidence likely to have a strong impact on the jury should be deemed "scientific" within the meaning of Frye and be subjected to its admissibility test.

While both courts used the correct test to determine whether expert testimony is "scientific," neither applied the test effectively. Both courts failed to conduct an adequate inquiry into the likely effects of the evidence on the jury. Such an inquiry would have first asked whether psychological research was available concerning the likely impact of the proffered evidence—or similar evidence—on the jury. While that inquiry alone would not have provided a definitive answer regarding jury impact, well-respected research studies also support the notion that jurors are not unduly impressed by, and

292. Taslitz, supra note 270, at 55. See also Steven M. Egesdal, Note, The Frye Doctrine and Relevancy Approach Controversy: An Empirical Evaluation, 74 Geo. L.J. 1769, 1774 n.26 (1986) (noting that advocates of the Frye doctrine are concerned about the weight juries will place on scientific testimony); Michael H. Graham, Expert Witness Testimony and the Federal Rules of Evidence: Insuring Adequate Assurance of Trustworthiness, 1986 U. Ill. L. Rev. 43, 53 (1986) ("[T]he test provides a preliminary screening to protect against a jury's natural inclination to give too much weight to scientific techniques . . . ."). Frye, properly understood, also involves the court in matters other than jury impact, such as the quality and quantity of experimental support for the technique used. See Taslitz, supra note 270, at 64. But before such inquiries can be undertaken a court must decide whether Frye should apply in the first place, for Frye does not control all expert evidence. See id. at 54. Whether the jury can properly evaluate the expert evidence is the most sensible test for making this decision. Id. at 55-56, 59-61.

293. See Taslitz, supra note 270, at 55. See also United States v. Addison, 498 F.2d 741, 744 (D.C. Cir. 1974) ("[S]cientific proof may in some instances assume a posture of mystic infallibility in the eyes of a jury of laymen . . . ."); United States v. Amaral, 488 F.2d 1148, 1152 (9th Cir. 1973) (noting that the Frye test is needed because of the "aura of special reliability and trustworthiness" that surrounds scientific testimony); People v. Kelly, 549 P.2d 1240, 1245 (Cal. 1976) ("Lay jurors tend to give considerable weight to 'scientific' evidence when presented by 'experts' with impressive credentials.").

294. See Taslitz, supra note 270, at 56.

can effectively evaluate, expert character testimony. Such studies validate Stoll’s holding and weaken Cavallo’s, but the Stoll court is still to be faulted for not even making the inquiry. Nor did either court inquire into historical, philosophical, or cultural sources to divine popular attitudes toward the proffered evidence. These types of inquiries can often prove useful in gauging the likely jury impact of certain evidence. Given that the charges involved in Cavallo and Stoll—rape and child sexual abuse—are subjects about which strong popular feelings and misconceptions exist, a more careful determination of jury impact may have been particularly beneficial.

The Cavallo court’s failure to evaluate evidence regarding jury impact was particularly egregious. That court failed to distinguish between “expert” testimony, to which Frye does not necessarily apply, and “scientific” testimony, to which Frye does apply. It concluded that merely labeling testimony “expert” might overawe the jury. The court offered no support for this extreme conclusion, a conclusion directly contrary to the way in which Frye is generally understood. Moreover, the Cavallo court’s broad condemnation of all psychological evidence as so “subjective” as to ensure a confusing “battle of experts,” was misplaced. Such generalizations are unwise because different types of evidence, even within a single class, may have very different effects. Therefore, the Cavallo court should have examined the particular category of psychological testimony involved to determine its likely jury impact.

The Stoll court did recognize the importance of this distinction among categories of psychological testimony. It warned that although the psychological evidence before the court would be un-

296. See infra text accompanying notes 632-640.
297. See Taslitz, supra note 270, at 22 & n.33 (explaining importance of popular culture as a guide to likely jury perceptions).
300. See supra notes 274-279 and accompanying text.
301. Cavallo, 443 A.2d at 1025. See also supra text accompanying note 277.
302. Professor McCord made a similar error but reached the opposite conclusion, arguing that all psychological testimony is unlikely to have an unfair jury impact because jurors will naturally be skeptical about psychological testimony. See McCord, Profiles, supra note 119, at 85-86 (“[M]ost jurors do not conceive of psychological research as very, if at all, 'scientific.' . . . In short, the jury most likely has the ability to fairly and intelligently weigh the strengths and weaknesses of psychological evidence without being overwhelmed or overawed by it.”).
likely to have an improper jury impact, other psychological evidence, notably the "post-hypnotic" testimony of a rape victim, may carry with it an "undeserved aura of certainty" requiring the application of Frye.\(^{303}\)

*Stoll*’s focus on the impact of a particular class of evidence on the jury is not, however, the case’s most interesting facet. The *Stoll* court also inquired into the manner in which the expert would testify and the quantity and quality of the information that the expert would provide in a particular case to determine whether the expert’s testimony would have an inappropriate impact. The court focused on the expert’s admissions that the MMPI had a seventy-percent accuracy rate, that it had proven to be wrong in at least one case in the expert’s own clinical practice, and that his approach was a “subjective, not strictly scientific” approach to the questions before him.\(^{304}\) Under these circumstances—specifically, when the expert disclosed weaknesses in the technique’s accuracy—those techniques seemed to be more art than science, subject to error, instead of being presented as a "foolproof" or "infallible truth."\(^{305}\) It is perhaps this aspect of *Stoll* that is most important, for it recognizes the crucial role that the law of evidence can play in maximizing the availability of information useful to the jury’s decision.\(^{306}\)

The question of jury impact can also be illuminated by contrasting *Cavallo* and *Stoll*, both of which involved attempts to disprove criminal acts, with the use of expert character testimony to prove or disprove a particular mental state. When mental state is at issue, psychological character testimony may sometimes be offered for the purpose of rebutting certain "myths" or preconceptions harbored by a jury.\(^{307}\) In such situations, there is less reason to fear that the

---

303. People v. Stoll, 783 P.2d 698, 710 (Cal. 1989). Indeed, the Supreme Court of California had earlier barred post-hypnotic testimony under *Frye*. See People v. Shirley, 723 P.2d 1354 (Cal. 1982). While some recent psychological research questions whether hypnotically refreshed testimony does indeed create an aura of certainty, see Vidmar & Schuller, *supra* note 135, at 169, it is at least arguable that “[t]he public has been barraged with misinformation concerning hypnosis and has resultantly been led to incorrectly believe that it provides a panacea for lost memory.” People v. Zayas, 546 N.E.2d 513, 516 (Ill. 1989).
304. See *Stoll*, 783 P.2d at 705, 712.
305. *Id.* at 712.
306. *See id.* (noting that the prosecution is free to rebut a defendant’s psychological character evidence through its own expert).
307. See Commonwealth v. Stonehouse, 555 A.2d 772, 785 (Pa. 1989) (“Thus, trial counsel was ineffective, and the absence of such expert testimony [on the battered woman syndrome] was prejudicial to appellant in that the jury was permitted, on the basis of unfounded myths, to assess appellant’s claim that she had a reasonable belief that she faced a life-threatening situation when she fired her gun at Welsh.”)
psychological evidence will "overawe" a jury, for the evidence is offered precisely to correct for the jury's potential misuse of other "evidence" arising either from jury ignorance or a jury's strongly, if irrationally, held beliefs.\textsuperscript{308}

The court must, of course, make a preliminary factual finding that a particular myth or preconception exists. It must also determine that the character testimony's impact in "correcting" for this myth or preconception is not outweighed by the "undue" impact of the character testimony.\textsuperscript{309} But if the court makes this finding, there is little justification for applying the \textit{Frye} standard, designed primarily to screen out potentially flawed expert evidence to which a jury might give undue weight. If the court has already made that finding, \textit{Frye} would serve no purpose.\textsuperscript{310} Moreover, even when no specific myth or preconception is proven, the jury's tendency toward stereotyping and shallow case logics suggests that any expert character testimony that focuses on what makes the defendant unusual or unique has a "correcting" function to play: forcing a jury toward the deep case logics required by the substantive law.

\begin{flushleft}
Cord, \textit{Profiles}, supra note 119, at 96 (discussing the "correcting" function served by non-traditional psychological evidence).
\textsuperscript{309} An interesting case on this point is People v. Leon, 263 Cal. Rptr. 77 (1989), in which a prosecutor in a child abuse case called a psychologist, who testified, among other things, that there are certain social myths regarding child sexual abuse. \textit{Id.} at 83-84. The court reversed the conviction and ordered a new trial partly because no evidence was offered under \textit{Frye} that the relevant scientific community generally agreed that these myths existed. \textit{See id.} at 84. The court did not address, however, whether testimony regarding "myths" was "scientific," thus requiring \textit{Frye}'s application. \textit{See id.} at 91-92 (Woods, J., dissenting, making this very point). The case is also interesting, however, for two other reasons.

First, the case highlights an important distinction: a court might, although the \textit{Leon} court did not, look to sources other than expert testimony (e.g., literature, philosophy, case law, history) to find that a myth exists and that testimony is necessary to respond to that myth, but the court can do so without alerting the jury to the myth's existence, thus eliminating \textit{Frye} concerns regarding "mythology." \textit{See Taslitz, supra note} 270, at 20-22, 123-25 (detailing role of mythology in the law of evidence).

Second, the prosecution's argument was offered to rebut myths that might lead a jury to believe that someone like the defendant could not commit the criminal \textit{act}. \textit{See Leon}, 263 Cal. Rptr. at 83-87. The case is unusual, however, because most expert character testimony regarding myths is offered when a defendant's state of mind is in question. \textit{See McCord, Profiles, supra} note 119, at 35-63 (reviewing many of the uses for nontraditional psychological evidence, the vast majority of which concern mental state); \textit{supra} text accompanying notes 129-135.

\textsuperscript{310} \textit{See supra} notes 292-293 and accompanying text.
\end{flushleft}
2. Is Expert Character Testimony “Novel”?—Although the Stoll court properly recognized that “scientific” evidence is governed by Frye only if the technique in question is “new” or “novel,” the court erred in seeing no real distinction between proving mental state or mental illness—“old,” previously accepted, uses of psychological evidence in California’s courts—but proving conduct—a logically “new” use. Justice Lucas, in dissent, made this point, emphasizing, without clearly explaining, that the use of objective personality tests “to formulate a psychiatric diagnosis of the defendant’s mental state is far different than using them to exclude defendant from the relevant class of offenders . . . .”

The Cavallo majority took a position similar to that of Justice Lucas, declaring that

the other areas in which psychiatric or psychological testimony is generally admitted in evidence are inapposite because such evidence there is admitted for different purposes . . . . Insanity or incapacity to stand trial, for example, represent conclusions about the state of mind of a person . . . . This is much different from asking experts to ascertain what characteristics non-rapists possess and to testify as to whether defendants have those characteristics.

While neither Justice Lucas in his Stoll dissent nor the Cavallo majority clearly explained why the difference should matter, matter it does. The novelty requirement, properly understood, serves to “bar relitigation of scientific reliability questions unless new evidence sheds a different light on the inquiry.” But scientific reliability does not exist in a vacuum; a scientific technique may be

311. See, e.g., People v. McCowan, 227 Cal. Rptr. 23, 30, (Cal. Ct. App. 1986) (holding that the trial court properly allowed “psychiatric testimony regarding [the defendant’s] mental condition” while excluding testimony on the defendant’s capacity to form the requisite mental state).


313. State v. Cavallo, 443 A.2d 1020, 1028-29 (N.J. 1982) (emphasis added). The court made a similar point earlier in its opinion in slightly different language: “[T]he evidence in this case purports to state the characteristics common to persons who commit certain acts, rather than suffer from certain illnesses . . . .” Id. at 1028.

314. See Taslitz, supra note 270, at 61 n.295. This purpose for the “novelty” requirement was advanced in light of criticism by Professor Starrs. See James E. Starrs, Frye v. United States Restructured and Revitalized: A Proposal to Amend Federal Evidence Rule 702, 26 Jurimetrics J. 252-53 (1986) (arguing that the “novelty” requirement could lead to inconsistent results by excluding novel, more reliable evidence while admitting older, less reliable evidence).
trustworthy for one purpose but not another.\textsuperscript{315} Consequently, even if a court has judged certain psychological techniques as valuable in examining a defendant's mental state, that does not necessarily mean that the techniques are equally valuable in predicting behavior. Psychological research supports the notion that behavior may often be controlled as much by situational factors as by one's essential "nature" or one's thoughts and feelings.\textsuperscript{316}

Moreover, even if the techniques used in \textit{Stoll} and \textit{Cavallo} were shown to be valuable in predicting some nonsexual behavior, that does not necessarily mean that the techniques would be of the same value in predicting violent sexual behavior.\textsuperscript{317} If, therefore, the evidence in either of the cases was "scientific," as the \textit{Cavallo} court incorrectly concluded, the novelty requirement demanded that \textit{Frye} be applied.

3. Are "Overt" and "Covert" Group Comparisons the Same?—A third significant matter in the two cases is that in \textit{Cavallo} there was an overt group comparison while in \textit{Stoll} the psychologist more heavily emphasized the clinical analysis of his patient as an individual. As noted earlier, the prosecutor in \textit{Stoll} indeed argued, albeit unsuccessfully on appeal, that the defense had to prove the general acceptance in the scientific community of a child molester "profile," and that "a person who does not fit the profile has not actually molested children."\textsuperscript{318} Only then, argued the prosecutor, could the defense psychiatrist testify that the defendant showed no "possibility for sexual deviancy in her personality profile."\textsuperscript{319} The prosecution apparently argued as well that such a profile had in effect entered into the psychologist's opinion and that such profiles make evidence

\textsuperscript{315} See, e.g., Edward J. Imwinkelried, \textit{The Methods of Attacking Scientific Evidence} \textsection 4-8 (1982) ("To satisfy \textit{Frye} your opponent must demonstrate general acceptance of the particular application of the instrument involved in your case."); Starrs, \textit{supra} note 314, at 258 ("Scientific validity for one purpose is not thereby automatically scientific validity for other, unrelated purposes.").

\textsuperscript{316} See \textit{supra} text accompanying notes 142-163. Indeed, common experience teaches us that our behavior does not always match our thoughts; thus a normally kind child in a school yard may feel sorry for the shy classmate being beaten by older children; but, despite feelings of compassion, the kind child may join in the beating to gain peer approval.

\textsuperscript{317} This is but one application of the concept of "external validity," that is, the need to inquire regarding the extent to which one can generalize the results of an experiment to other situations. \textit{See generally} David W. Martin, \textit{Doing Psychology Experiments} 118 (3d ed. 1991).

\textsuperscript{318} People v. Stoll, 783 P.2d 698, 706 (Cal. 1989).

\textsuperscript{319} \textit{Id.}
"seem" scientific to a jury, triggering the *Frye* test.\textsuperscript{320} The appellate court, however, rejected the prosecutor's reading of the record of the psychologist's testimony, concluding instead that "no psychological 'profile' entered into his diagnosis."\textsuperscript{321} While the court, therefore, rejected the claim that there was an explicit group comparison in the form of a profile that would overawe the jury, the court nevertheless noted in dicta, "[w]e are not persuaded that juries are incapable of evaluating properly presented references to psychological 'profiles' and 'syndromes.'"\textsuperscript{322}

The *Cavallo* court, in contrast, repeatedly stressed that one of the questions before it was "whether psychiatrists agree that rapists share particular mental characteristics."\textsuperscript{323} The assumption that they do, of course, was implicit in the proffered testimony that rapists are aggressive, violent people—a "mold" into which the psychiatrist would testify the defendant did not fit.\textsuperscript{324} The court worried that this explicit reference to a rapist "mold" would turn the issue in the jury's mind from whether the defendant committed the crime to whether he manifested the "characteristics of a 'rapist.'"\textsuperscript{325}

Both courts were wrong, albeit for slightly different reasons. As discussed above,\textsuperscript{326} most expert character evidence requires comparison to a group,\textit{even if that comparison is implicit rather than explicit}. Thus, the expert's opinion in *Stoll* was based partly on the MMPI, which in turn was based on experience with the behavior of groups of people receiving particular scores or score patterns on the test. Moreover, to argue, as the defense did in *Stoll*, that the defendant had essentially a "normal" personality function and was, therefore, less likely to commit a sex crime than those who are not "normal," involves at least two group comparisons: to "normal" people (who are not likely to commit the crime) and to "abnormal" people (who are more likely to commit the crime). The logical process that the jury should undertake in *Stoll* was thus, at least partly the same as in

\textsuperscript{320} The dissent better articulated the prosecution's apparent position, noting that testimony regarding the MMPI "validity scales" supposedly assured that no intentional concealment of deviancy could occur, thus making the technique appear to be "a 'personality print,' i.e., a supposedly foolproof scientific method of matching a suspect's personality traits with the offense in question, much in the same way that voiceprints, fingerprints, blood or semen samples, are used to match or exclude a suspect." \textit{Id.} at 718-19 (Lucas, J., dissenting).
\textsuperscript{321} \textit{Id.} at 708.
\textsuperscript{322} \textit{Id.} at 714 n.22.
\textsuperscript{323} State v. Cavallo, 443 A.2d 1020, 1027 (N.J. 1982).
\textsuperscript{324} See \textit{id.} at 1021-22.
\textsuperscript{325} \textit{Id.} at 1025.
\textsuperscript{326} See supra text accompanying notes 121-135.
Cavallo: to determine whether members of the group of child sex abusers share some common features that identify them as more likely to commit a sex crime than is true for members of the group of those who do not abuse children.

This does not mean, however, that there is no distinction between the two cases. The overt use of probabilistic evidence for group comparison arguably might create precisely the kind of impression of scientific precision with which Frye is concerned. For example, testimony that seventy percent of all child sex abusers receive a particular score on the MMPI might so impress jurors that they will assume that the defendant cannot be guilty of child sex abuse if his test does not reflect the necessary score. This fear that juries will be highly impressed with numbers and yet not really understand their significance has indeed been a concern of many courts and evidence scholars.\textsuperscript{327} But this concern, voiced in the Cavallo opinion, is unwarranted because the empirical research on jury behavior strongly suggests that, at least under certain conditions, juries do not overvalue explicitly mathematical or probabilistic testimony; indeed, they may sometimes undervalue it.\textsuperscript{328} While the Stoll court was therefore wrong to conclude that profiles and more "individualized" clinical assessments are fundamentally different—they are not\textsuperscript{329}—the court's dicta expressing faith in the jury's ability to handle profiles is probably well taken. It should be remembered, however, that even if juries may generally be capable of handling profile testimony, a trial judge should nevertheless conduct a particularized assessment of the impact of the type of profile evidence offered and of the manner in which that evidence is offered in each case.\textsuperscript{330}

\textsuperscript{327} See, e.g., Laurence H. Tribe, Trial by Mathematics: Precision and Ritual in the Legal Process, 84 HARV. L. REV. 1329 (1971); infra note 328.

\textsuperscript{328} David McCord, A Primer for the Nonmathematically Inclined on Mathematical Evidence in Criminal Cases, 47 WASH. & LEE L. REV. 741 (1990) [hereinafter McCord, Nonmathematically Inclined] (summarizing much of the research in this area and specifying the conditions under which statistical evidence is unlikely to unduly influence the jury); David Faigman & A.J. Baglioni, Bayes' Theorem in the Trial Process: Instructing Jurors on the Value of Statistical Evidence, 12 L. & HUM. BEHAV. 1, 16 (1988) ("[O]ur results . . . suggest, contrary to Tribe's assertion, that an expert's Bayesian formulation will not overwhelm the average trier of fact. Courts, it seems, should be less concerned with jurors being overwhelmed by the complexity of statistical techniques and more concerned with impressing upon jurors the relevance of those techniques.").

\textsuperscript{329} There is, however, one important distinction between profiles and more individualized assessments: the latter turn to a greater degree on the clinical "arts" than on science. See infra text accompanying notes 546-573. This distinction is, however, more important for purposes of Federal Rules of Evidence 403 and 702 than for Frye. See id. 330. See supra text accompanying notes 293-309.
4. Is the "Cathartic Function" Served by Frye?—Despite the emphasis thus far on a case-by-case judgment of whether Frye should apply to psychological character evidence testimony, there is a strong argument that the Frye standard should never govern expert character evidence offered by the defense in criminal trials. This argument flows from an understanding that the law of evidence serves goals other than truth-seeking, among which is what Professor Leonard has called the "cathartic function."331

Understanding the "cathartic function" begins with a recognition that, in addition to being a search for "the truth," trials serve to give parties a formal opportunity to vent their feelings before a group of average citizens able to understand the justice of the parties' positions; trials also serve to satisfy the community that order prevails and that justice is being done.332 "‘Catharsis,’ in the context of trials can therefore be defined as a vital point of satisfaction that must be experienced both by the litigants in any particular case and by society in general as it develops its impressions about the processes of the law."333 Professor Leonard argues that character evidence serves this cathartic function.334 Consequently, he proposes a balancing approach for character evidence in civil cases, rather than a per se bar, so that a trial judge can weigh the dangers of the testimony against its benefits, including the testimony's value in serving the cathartic function.335

In criminal cases, Professor Leonard would go further and grant defendants an absolute right to offer character evidence, apparently on a theory of the overriding importance of the cathartic function in criminal cases.336 Professor Leonard offers little justification for this theory. He does discuss, however, the fact that the prosecution must not have the same right as the defense to free admissibility of character evidence.337 This is because the cathartic value of allowing the prosecution to offer evidence regarding a de-

332. Id. at 39.
333. Id. Professor Leonard thus redefines the notion of truth-seeking in a trial: "[W]e also share an idea of truth as representing a point of intuitive, even emotional satisfaction with results reached by the trial." Id. at 32.
334. Professor Leonard does not fully explain why this is so, other than to argue that, given the often-low probative value of lay character testimony, only catharsis can explain why we still, at times, permit character evidence. See id. at 54-61.
335. See id.
336. See id. at 56-58.
337. See id. at 50 n.254.
fendant's character is limited in that (1) society's sense of loss from a single guilty person's going free is not as great or as immediate as that suffered by a convicted defendant who is not permitted to offer good character evidence, and (2) there is a significant chance that a jury hearing evidence of "bad" character will be unduly moved by passion and prejudice, an outcome less likely when a jury hears "good" character evidence from the defense.\(^{388}\)

Social psychological theory provides strong support both for the concept of a cathartic function and for the argument that this cathartic function has an especially strong role to play in criminal cases. First, a long line of research suggests that when disputants have greater control over the procedures by which they are judged, including the admissibility of evidence, they perceive the verdicts as fairer and are therefore more accepting of even unfavorable verdicts.\(^ {389}\) Indeed, the perceived fairness of the procedures is a significant factor increasing the likelihood that parties involved in a dispute will obey the law in the future.\(^ {390}\) The perceived fairness of procedures also improves society’s satisfaction with the legal system generally and public support for the courts specifically.\(^ {391}\) The cathartic function may therefore play at least some role in reducing

\(^{388}\) See id. While Professor Leonard would, once a defendant has “opened the door,” permit the prosecution to offer character evidence, he would allow this solely for impeachment purposes. He deems this limitation essential to “protect the defendant's basic right to be tried according to the facts of the case and not according to her general character.” Id. at 57.


\(^{390}\) See id. at 76-81.

\(^{391}\) See id. at 64-76. But cf. Alan Hyde, The Concept of Legitimation in the Sociology of Law, 1983 Wis. L. Rev. 379. Professor Hyde argued that people obey the law primarily because of a fear of punishment and other factors, but not because the law promotes “legitimacy.” Id. at 385-86, 426 (More precisely, he argued that there is inadequate empirical evidence that legitimacy plays a role.). Without engaging in an extended critique of Professor Hyde's argument or the meaning of legitimacy, it is sufficient to note this: psychological research published since Hyde’s article suggests that the perception of fair procedures is indeed a factor promoting compliance with, and public support for, the law. See TOM R. TYLER, WHY PEOPLE OBEY THE LAW 42-56, 62-68, 178 (1990) (finding a connection among perceptions of fair procedures, legitimacy, and compliance with law). Indeed, the importance of procedural justice grows as the stakes grow. See id. at 105 (discussing empirical data regarding the importance of fair procedures in felony trials). A critical question in promoting respect for and compliance with the law is, therefore, whether citizens believe they have been, or will be, treated fairly by the criminal justice system. See id. at 104-12 (discussing the relationship between experience with the authorities and public support for the law).
both civil and criminal wrongs and improving respect for both the civil and criminal justice systems.

Second, the substantive rules of criminal law generally reflect society's highest judgments of morality. Procedures that are inconsistent with these judgments are likely to be perceived as particularly unfair, thus magnifying the individual and social impact of procedural injustice. In particular, when criminal punishment is imposed on the basis of an individualized inquiry into a defendant's thoughts, feelings, or actions, that defendant will particularly resent a conviction if denied an opportunity to present her version. Correspondingly, society has less reason to be comfortable that it is visiting its moral judgments only upon the truly guilty when those judgments demand individualized justice, but the accused is offered no opportunity to argue based upon the full uniqueness of her situation. Rephrased, the cathartic function is disserved when rules of evidence bar testimony that would move the jury toward the deep case logics demanded both by popular morality and substantive rules of criminal law, leaving a defendant to be judged by shallow case logics inconsistent with an individualized approach.

What does any of this have to do with the Frye rule? The answer is that in certain contexts, as Professor Leonard wisely argues, there must be a balancing approach that weighs the "truth-seeking function" of character evidence against its cathartic function. Because Professor Leonard addresses expert character testimony only in passing, he does not address the implications of Frye in any depth. But the implications are clear: if the Frye rule applies to

342. See supra Part I.C.

343. In criminal cases, as noted above, Professor Leonard argues for a per se rule of admissibility for defense-proffered lay character testimony. See supra text accompanying notes 336-338. Whatever sense that rule may have for lay testimony—a debatable point—it makes no sense for expert testimony because there would then be no safeguards against admitting confusing, unreliable, and misleading expert evidence. See infra text accompanying notes 642-652.

344. Professor Leonard states in a footnote, however, that if Frye applied to expert character evidence, the appropriate inquiry should not be whether the techniques used are generally accepted in the relevant scientific community but rather, whether that community generally accepts that expert character testimony is more reliable than lay character testimony. See Leonard, supra note 331, at 48 n.249. He later suggests that expert character evidence would indeed be more reliable than lay character testimony. See id. at 52. This Article will argue in greater depth that expert character testimony is certainly more "helpful" (a term not identical with "reliable") than similar lay testimony. However, it is not clear that if Frye were applied in the way Professor Leonard suggests, the relevant scientific community would generally agree that expert character testimony is more "reliable" than lay character testimony. See, e.g., Richard E. Nisbett & Lee Ross, Human Inference: Strategies and Shortcomings of Social Judgment 93-101 (1980)
expert character testimony, there is no opportunity to weigh catharsis against truth, for techniques that are not generally accepted will automatically be excluded from evidence. This per se bar is unwise in criminal cases where the cathartic function deserves special attention.345

The central point is that neither the impact on truth nor the impact on the cathartic function can be considered if Frye applies to expert character evidence offered by the defense. For the reasons articulated by Professor Leonard, however, the cathartic function has less force when applied to evidence offered by the prosecution and the primacy of the truth-seeking function may therefore demand Frye's application to prosecution-generated expert character testimony despite the resulting asymmetry.346

C. Pragmatic Relevancy and Psychological Character Testimony

The primary alternative to Frye—the "pragmatic relevancy approach"347—would permit a broader truth-related inquiry and a weighing of truth-related values against process-related ones ("cathartic" values) that Frye prohibits.348 This alternative approach treats scientific evidence like all other evidence—and all evidence must survive Federal Rule of Evidence 403's balancing of probative

(summarizing some similar inferential errors regarding human behavior that are often made by both lay persons and clinicians). But see infra text accompanying notes 385-405, 575-616 (arguing for ways to improve expert "reliability" and for the superior value of expert character evidence to the jury).

345. Permitting a defendant to decide whether to offer expert character evidence may advance truth as well as catharsis; giving a defendant this control increases the likelihood that she will produce evidence of individualized circumstances, thus helping to overcome jury tendencies toward trial by stereotype. See Lind & Tyler, supra note 339, at 38 (noting that parties to a dispute are more likely to produce individualized information than uninvolved decision-makers).

346. See supra notes 336-338 and accompanying text.

347. See Paul C. Giannelli, The Admissibility of Novel Scientific Evidence: Frye v. United States, a Half-Century Later, 80 Colum. L. Rev. 1197, 1235-39 (1980). Professor Giannelli and other commentators have used the term "relevancy," rather than "pragmatic relevancy," to describe this approach. However, as discussed earlier, the term "pragmatic relevancy," coined by Professors Mueller and Kirkpatrick, is clearer. It distinguishes the pragmatic balancing of the favorable and unfavorable qualities of evidence that are at the heart of this approach from the narrower inquiry that characterizes simple "logical relevancy." See supra text accompanying notes 13-14.

348. See Leonard, supra note 331, at 53-60 (explaining that the "guided discretion" accorded to the court under Federal Rule of Evidence 403—the source of authority for the pragmatic relevancy approach, see infra note 349—includes the discretion to balance cathartic against truth-related values).
value against countervailing considerations. Various commentators have identified an array of factors that should be considered in weighing this balance. This Article focuses on three broad factors identified by Professor McCord as specifically useful in engaging in relevancy-balancing for “non-traditional” psychological evidence: reliability, necessity, and understandability. Reliability is concerned with the accuracy of the technique, that is, whether it correctly diagnoses what it purports to diagnose (corresponding to the scientist’s concept of “validity”) and how consistent the technique is, that is, how often experts in the field agree on the diagnosis (corresponding to the scientist’s definition of “reliability”). Necessity asks whether the testimony is necessary in the context of the case, and, if so, how necessary? Understandability asks whether jurors can understand the testimony and give it proper weight.

1. Reliability.—Reliability requires an inquiry into psychological character evidence’s accuracy and consistency as a predictor of thoughts and behavior and the qualifications of the experts who test-

349. See PAUL C. GIANNELLI & EDWARD J. IMWINKELRIED, SCIENTIFIC EVIDENCE 31 (1986) (“The principal alternative approach to the Frye test is to treat scientific evidence in the same way as other evidence, weighing its probative value against countervailing dangers and considerations.”) (emphasis added)). One source of this weighing is Federal Rule of Evidence 403: “Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.” FED. R. EVID. 403. A second source is Rule 702, which permits an expert to testify only if “scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue . . . .” FED. R. EVID. 702 (emphasis added). Obviously, evidence cannot “assist the trier of fact” if the evidence’s probative value is outweighed by prejudice, confusion, or the like.


351. McCord, Profiles, supra note 119, at 94-101. Professor McCord labels his approach a “will assist the jury” approach but recognizes that it is but a variant of the “relevancy” approach modified for the “softer” science involved in nontraditional psychological evidence. Id. at 93-94.

352. Id. at 95.

353. Id. at 98-99.

354. Id. at 100. Professor McCord identifies one additional key factor: “importance,” which asks how crucial to the resolution of the case is the testimony being offered. Id. at 101. This Article concludes that the “importance” factor is not helpful in understanding expert character testimony. See infra text accompanying notes 642-667.
The former inquiry requires a comparison between the relative worth of experimental data and clinical judgment as bases for character opinions and a separate examination of the probative value of “negative inferences,” arguments of relevancy based on proving the absence of a particular trait.

a. The Predictive Power of Traits.—To assess the probative value of expert psychological evidence—its tendency to make a fact of consequence more probably true than would be so without the evidence—requires an evaluation of how well the “personality trait” concept predicts human thoughts and behaviors. Two major challenges have been made to the predictive power of traits. First, the correlation between any specific measure of a trait and the behavior being predicted is generally quite low. This suggests that “traits”—roughly defined as stable cross-situational individual predispositions—have little real existence and little predictive value. Second, even the staunchest defenders of the idea that traits do exist and can be useful concede that psychologists cannot predict any individual instance of behavior with any degree of confidence based on knowledge of the “trait” in question.

355. See McCord, Profiles, supra note 119, at 98-99. Professor McCord also identified general acceptance as a relevant factor in determining the reliability of nontraditional psychological evidence. Id. at 99. It is important to remember, however, that under a relevancy approach, general acceptance is merely a factor and is not determinative of the admissibility question, in sharp contrast to the Frye approach. Because the acceptance level of a technique involves a factual determination to be made as to each particular technique—and this Article does not focus on any single such technique—no further discussion of this factor is necessary here.

356. See MCCORMICK ET AL., supra note 16, § 185, at 541 (defining probative value).

357. See Giannelli, supra note 347, at 1235-39 (“The probative value of scientific evidence... is connected inextricably to its reliability.”). The inquiry as to the degree of probative value is, therefore, also related to the question: “How good a predictor is this evidence?”

358. See infra text accompanying notes 363-374.

359. See supra text accompanying note 142.

360. See Susan M. Davies, Evidence of Character to Prove Conduct: A Reassessment of Relevancy, 27 CRIM. L. BULL. 504, 514 (1991) (declaring that critics of trait theory “effectively denied the existence of stable personality traits...”); Teree E. Foster, Rule 609(a) in the Civil Context: A Recommendation for Reform, 57 FORDHAM L. REV. 1, 27-37 (1988) (stating that a majority of psychologists agree that the “trait-oriented approach” does not shape future behavior); Leonard, supra note 331, at 25-31 (concluding that studies regarding the poor predictive power of character traits demonstrate that much character evidence entirely lacks probative value).

361. See Davies, supra note 360, at 517 (“Contemporary trait theorists willingly concede that they are unable to predict a single instance of behavior on a particular occasion with confidence.”).
The first of these challenges is based, however, on the failure to "aggregate," a methodological flaw in measuring traits. The second challenge ignores the important truth that knowledge of "average" behavior can be highly relevant, even though averages do not permit precise predictions of individual instances of behavior. Moreover, both challenges ignore the various ways in which the predictive power of trait-related concepts can be improved or supplemented.

i. The Principle of Aggregation.—The principle of aggregation, ignored by many critics of trait theory, can be understood only by considering the meaning and nature of "correlation" and of "correlation coefficients." Correlation describes "the degree to which two sets of measures, such as intelligence and academic achievement, are related." The most common type of correlation coefficient is a numerical index ranging from minus 1—a perfect inverse relationship—to plus 1—a perfect direct relationship. The average correlation between different purported behavioral measures of the same personality trait is typically in a range between .1 and .2. In lay terms, this means that there is almost no gain in accuracy of prediction about a later situation by virtue of knowing how someone behaved in an earlier situation.

Correlations between scores on paper-and-pencil measures of personality ("personality tests") and individual observations of a behavior demonstrating the existence of the measured trait (for example, between measuring high on a "friendliness" score on a personality test and actually being observed as friendly in a particular situation) do little better, rarely exceeding the .2 to .3 range. Indeed, virtually no coefficients, either between individual pairs of

362. See infra text accompanying notes 363-384; Davies, supra note 360, at 513-18 (1991) (discussing the usefulness of traits to accurately predict average behavior and noting that relevancy does not equate to sufficiency of evidence); Seymour Epstein, The Stability of Behavior: I. On Predicting Most of the People Much of the Time, 37 J. PERSONALITY & SOC. PSYCHOL. 1097, 1102 (1979) [hereinafter Epstein, Most People] ("A trait is a generalized tendency for a person to behave in a certain manner over a sufficient sample of events and does not imply that he or she will exhibit trait-relevant behavior in all situations or even on all occasions in the same situation.").
363. Aiken, TESTING, supra note 188, at 26.
364. Id.
366. For example, a .16 correlation between being friendly in situation 1 and proving indeed to be friendly in situation 2 increases the likelihood of friendliness in the latter situation to only 55%, compared to a 50% likelihood absent a correlation study but given only two choices: friendly or unfriendly. Id.
behavioral measures or between a personality test and an individual behavioral measure, exceed .3.\textsuperscript{367} This data suggests to some that cross-situational traits either do not exist or are not a useful concept.\textsuperscript{368}

But these correlations are misleading because, as Epstein demonstrated, they overlook the benefits of "aggregation."\textsuperscript{369} Individual responses, like single items on a test of any kind, are usually highly unreliable and reflect the impact of many systematic and random factors other than the one being studied.\textsuperscript{370} Thus even an excellent student may get individual exam questions wrong or have an off day and get a poor grade on a particular test.\textsuperscript{371} But what marks the good student apart from lesser students is that, on average, the good student scores much better than poorer students.\textsuperscript{372} Averages in turn are obtained by aggregating scores, that is, combining multiple measures of scholastic ability and averaging them.\textsuperscript{373} Similarly, if traits exist, the average of multiple behavioral observations should lead to much larger correlation coefficients. Indeed, this has proven to be the case. For example, if twenty-five independent behavioral measures of a subject's extroversion are averaged and that average correlated with the average of another twenty-five independent behavioral observations of the subject's extroversion, the correlation coefficient rises to .83, a very impressive coefficient.\textsuperscript{374}

\textit{ii. Average Behavior.—}While aggregation demonstrates, therefore, that cross-situational dispositions do exist,\textsuperscript{375} it must be remembered that aggregation is required precisely because there is much variability in individual instances of behavior for each person.

\textsuperscript{367} See id.

\textsuperscript{368} Id. at 95-96 (noting that Walter Mischel, one of the first psychologists to summarize the literature in this area and point out its significance, "challenged us to entertain the possibility that those low correlations might be capturing an important truth about human behavior, that is, that cross-situational consistency might be the exception and behavioral specificity the rule").


\textsuperscript{370} See Ross & Nisbett, supra note 365, at 107.

\textsuperscript{371} Id.

\textsuperscript{372} Id.

\textsuperscript{373} Id.

\textsuperscript{374} See id. at 107-08.

\textsuperscript{375} See Epstein, Stability, supra note 369, at 268 ("The following statements are all true: (1) behavior is situationally specific, (2) behavior is general across situations, and (3) people have broad cross-situational response dispositions." (emphasis added)).
Indeed, what modern social psychologists have realized is that all individuals will show a wide range of responses and all will make responses close to the overall population average more often than they will make responses that are extreme. This variability and clustering around the population average means that it is difficult to predict behavior in a particular situation based upon a cross-situational trait derived from aggregation, just as it is difficult to predict behavior based on only an individual, "unaggregated" score.

But this does not mean that correlation coefficients derived from aggregating, or even those not so derived, are logically irrelevant if the concept of "relative likelihood" is understood. Assume, for example, based on a single, randomly sampled measure for each of two individuals, Paula and Jane, that the likelihood of Paula's engaging in a particular extreme behavior, for example, extroversion, is 4.5 percent, while the likelihood of Jane's doing so is an even smaller .9 percent. While the extreme behavior is thus highly unlikely for both individuals, Paula is nevertheless five times more likely than Jane (5 x .9 = 4.5) to display the extreme behavior.

This difference in "relative likelihood" is unquestionably "relevant." Evidence is relevant if it makes a fact of consequence more or less probable than would be true without the evidence. In an assault case, knowing that Paula is five times more likely than Jane to engage in violent behavior, a variation on the above example, obviously makes it more likely than otherwise that, as between Paula and Jane, Paula is the one who committed the assault.

Similarly, the relative likelihood of Paula's being violent might be compared to the likelihood of a "typical" person's doing so, that is, a person who scores precisely at the population mean, thus informing the jury that Paula is X number of times more—or less—likely to be violent than is the average person. A related concept is comparing Paula's rate of violence with the population base rates, that is, the proportion of people in the general population who manifest the behavior of interest. For example, comparison of Paula's rate of violence with those who display a smaller rate of violence, and those who display a higher rate, can place Paula's relative

376. See Ross & Nisbett, supra note 365, at 110-11.
377. See id. at 107-15.
378. Id. at 116.
379. See id.
381. Aiken, Testing, supra note 188, at 110.
violence in the context of the general population.\textsuperscript{382}

Moreover, if we aggregate data (which the above examples did not), we enhance the statements of relative likelihood dramatically, so that we might be able to improve the confidence of our estimates in the cases of Paula and Jane to say that Paula is 100 times more likely than Jane to be violent.\textsuperscript{383} Additionally, these same relative likelihood data can be mathematically manipulated to enable us to use past extreme behavior to calculate the probability of an individual’s engaging in that same behavior across a certain specified number of future observations. Thus, if Paula scored in the ninety-eighth percentile on the basis of a single past observation of a behavior, mathematical formulae enable us to estimate that there will be a thirty-four percent chance that she will engage in another extreme behavior, in the top two percent of the population, at least once in the next ten observations.\textsuperscript{384} This, too, helps to increase the likelihood that she committed an assault if that extreme behavior is assaultive violence.

iii. \textit{Improving Predictive Power}.—Several further observations must be made about ways to increase the predictive power of expert character evidence. First, as noted earlier, virtually all psychologists now agree that much of behavior is accounted for by situational factors, not merely by an individual’s different traits.\textsuperscript{385} Thus, traits defined in more situationally specific terms are more likely to be accurate predictors of behavior.\textsuperscript{386} In any event, even if broader cross-situational traits are used, predictions cannot be of much value without consideration of the likely impact of the situation, par-

\textsuperscript{382} See Ross \& Nisbett, \textit{supra} note 365, at 138-39 (concluding that cross-situational trait-based predictions can be made with good prospects for accuracy if certain circumstances exist, including: (1) the prediction is based on a large, diverse sample of past observations, (2) the prediction addresses the relative likelihood of extreme outcomes or events, when the actors have in the past shown themselves to be extreme relative to others, and (3) the predictions are mindful of population base rates). Ross and Nisbett further note that when these conditions are met, “trait information can in theory predict even to novel situations,” \textit{id.} at 139, an important point given that many crimes are indeed first-time “novel” experiences as, for example, when a wife kills her husband.

\textsuperscript{383} See \textit{id}. at 117.

\textsuperscript{384} See \textit{id}.

\textsuperscript{385} See Davies, \textit{supra} note 360, at 518 (“The trait-versus-situation controversy in the field of personality and social psychology has produced widespread agreement that behavior is simultaneously a function of disposition and situation, and their mutual interaction.”).

\textsuperscript{386} See, e.g., Ross \& Nisbett, \textit{supra} note 365, at 138 (noting that individual difference information can be invaluable “where one has base rate information for the particular individual for the particular situation. Thus your prediction that Jack will talk a lot at the lunch table today, as in the past, is bankable.”).
particularly when there is supporting experimental data as to the performance of people under certain situations.\textsuperscript{387}

For expert character testimony to be most useful, therefore, a witness must discuss not only the subject’s relevant “traits,” but the likely impact of the situation at issue, and the interaction between those traits and those situations. Only then can a fact-finder truly come to understand a defendant’s thoughts and behavior. This may need to be a wide-ranging inquiry, because many factors beyond the most obvious ones play into the “situation,” including the social roles played by the participants, their cultural background, and the nature of the “audience” before whom a defendant is “playing.”\textsuperscript{388}

Second, even given a particular trait and a particular situation, a person may react differently depending on how that person subjectively construes the situation.\textsuperscript{389} For example, two subjects might share the trait of displaying anger toward those who do not like them. However, one subject might perceive person A as disliking that subject, leading to anger, while the other subject may construe the same behavior by A as showing that A indeed likes the subject, thus avoiding anger. But how people construe situations often differs dramatically and even the same individual may display unpredictable changes in construal.\textsuperscript{390} On the other hand, differences in construal may themselves “reflect longer term, more idiosyncratic differences in personal history and temperament.”\textsuperscript{391} Studying “idiosyncratic” differences and unique ways of construing a situation in turn requires a more “idiographic” approach to the study of personality,\textsuperscript{392} an approach that demands some unavoidable measure of clinical judgment. Judgments based at least in part on clinician observations will be necessary because the idiographic approach requires a willingness to observe the individual actor in a substantial

\textsuperscript{387} See Davies, supra note 360, at 519 (“The ultimate question to be resolved by the forensic fact finder is most analogous to that of concern to the interactionist, that is, how this person (for example, a witness) would behave in a particular situation (such as, when giving testimony at this trial).”).

\textsuperscript{388} See Ross & Nisbett, supra note 365, at 150-54, 156-58, 169-88 (discussing the various factors in the predictability of behavior).

\textsuperscript{389} See id. at 67-89, 165-66 (discussing the concept of construal and establishing its importance in predicting behavior).

\textsuperscript{390} Id. at 68-69.

\textsuperscript{391} Id. at 165.

\textsuperscript{392} See id. at 163 (“The approach that is called for . . . must be idiographic in spirit . . . . That is, we will need to know different things about different people in order to appreciate the distinctiveness and coherence in their behavior and in order to anticipate when, and how, their behavior will prove consistent and predictable.” (emphasis added)). See also supra note 164 (defining and contrasting “idiographic” and “nomothetic” approaches to personality).
sample of situations. It also requires a review of the actor’s personal history, needs, goals, and the like, in an attempt to divine behavioral consistencies. In this way the particular and unique set of situations in which the actor will manifest her particular dispositions can be anticipated.

Such an idiographic approach also requires that attention be paid to individual goals and preferences, for understanding what they are and how important they are to a particular person may explain the true consistency in that person’s behavior. Competencies and capacities, that is, abilities to achieve personal goals and satisfy personal standards, are also important. One person may, for example, know how to get along with peers while another does not. Differences in attributional styles and perceptions of self-efficacy—that is, whether individuals feel in control of their life and responsible for their happiness, or instead feel like helpless pawns—can dramatically affect behavior.

Furthermore, people have different “self-schematas,” that is, different generalized understandings of the self. Some people substantially organize their behavior around the concept of dependence, others around the notion of independence, and still others consider either concept irrelevant. Those who are schematic for dependence or independence will provide “more evidence when asked to defend assertions that they are dependent or independent, and they are highly resistant to information that seems to contradict their self-schemas.” A related concept, “self-monitoring,” reflects the notion that different people monitor themselves along different dimensions in different situations. But there are also differences in the extent to which people engage in self-monitoring

393. See Ross & Nisbett, supra note 365, at 103.
394. See id.
395. See id.
396. See id. at 163-64. One preference that illustrates the important interaction between personal goals and situations is our preference to place ourselves in particular situations. While the situation may be a powerful determinant of our behavior, there may be underlying consistencies regarding the situations in which we choose to place ourselves, and thus regarding the behavior that we prefer those situations to elicit from us. See id. at 154-56.
397. See id. at 164-65 (noting that competence and capacity may predict a person’s ability to achieve goals).
398. See id. at 166 (noting that the conception of the self guides behavior and interprets one’s behavior toward others).
399. Id.
400. Id.
401. Id.
402. Id. at 167.
of their behavior or its impact on others that can be measured with a pencil-and-paper test.\textsuperscript{403}

Finally, it should be noted that, while the presence of mental "illness" or the existence of a personality "disorder" is not necessary to validate expert character opinions, such diagnoses may improve the predictive power of character evidence for a very simple reason: disorders by definition involve a certain "rigidity" in thinking or behavior, an unwillingness or inability to think or act differently in different circumstances.\textsuperscript{404} "Rigid" behaviors are by definition quite consistent, thus improving the confidence with which a clinician can predict the patient's behavior.\textsuperscript{405}

b. Experimental Data versus Clinical Judgment.—While individual differences in the predisposition to think or act in a certain way are real and should, in theory, be measurable, the value of a mental health expert's opinion in a particular case turns on more than this truth. It turns, in addition to matters already discussed, on two further factors: (1) the adequacy of any supporting experimental or quasi-experimental data; and (2) the value of "clinical judgment."

i. Experimental Data.—Experimental data is most important at the group level of analysis. As an illustration, assume that a defendant, a child brutally abused by his father, argues that abused children become abnormally fearful of their abusing parent. The defendant claims that, because he was such a child, he honestly believed that his father's actions endangered his life; he therefore killed his father in what he believed to be self-defense, even though the killing was prearranged.\textsuperscript{406} But the validity of this argument should turn at least partly on whether there is experimental data to support the defendant's claim that such an abnormal fear exists among abused children.

That inquiry may reveal that the claim has never been subjected to experimental challenge. More interestingly, the inquiry may in-

\textsuperscript{403} See id. (noting that the pencil-and-paper test to measure self-monitoring reveals differences in self-monitoring across many behavioral dimensions and social contexts).

\textsuperscript{404} Interview with Ellen McDaniel, M.D., a well-known forensic psychiatrist, in Washington, D.C. (July 26, 1991) (describing the reduced plasticity of behavior in differing situations as a distinguishing characteristic of personality and other psychological disorders); IRA D. TURKAT, THE PERSONALITY DISORDERS: A PSYCHOLOGICAL APPROACH TO CLINICAL MANAGEMENT 11-12 (1990) (noting that DSM-III-R classifies personality traits as "disorders" only when the traits are inflexible and maladaptive and "cause either significant impairment in social or occupational functioning or subjective distress").

\textsuperscript{405} McDaniel interview, supra note 404.

\textsuperscript{406} This hypothetical is based on Jahnke v. State, 682 P.2d 991, 995-96 (Wyo. 1984).
stead reveal that existing data is inadequate. Judging the adequacy of experimental research is not difficult for lawyers, particularly when the lawyer is aided by a qualified expert witness.407 Professor Faigman has indeed identified five common errors made by social scientists that can easily be watched for by courts and attorneys.

First, samples are often unrepresentative of the larger group about which conclusions are to be drawn, or are too small to justify such conclusions.408 For example, in studying jury behavior, it would be unwise to generalize based upon a small sample of college sophomores, whose reactions may not represent those of the broader population. Another common concern about the fairness of generally applying the results of a study409 is that laboratory studies may ignore some factors at play in the real world and may thus require confirmation by field studies, even though such studies raise their own problems of "internal validity."410

407. See David L. Faigman, To Have and Have Not: Assessing the Value of Social Science to the Law as Science and Policy, 38 EMORY L.J. 1005, 1080 (1989) [hereinafter Faigman, Social Science and Policy] ("Even the most difficult concepts used by social scientists are no more difficult than the more esoteric concepts lawyers employ regularly. The proper question is whether lawmakers have the inclination or motivation to devote the time necessary to learn the methods of social science."). See generally McCord, Child Abuse, supra note 295 (illustrating a lawyer's effective use of social science research to analyze a difficult legal issue).

408. See Faigman, Social Science and Policy, supra note 407, at 1059 (noting that small samples decrease the chances that the observed behavior represents the behaviors of the large and complex population and that unrepresentative samples bear little resemblance to the actual population of interest).

409. This is what scientists call "external validity." See MARTIN, supra note 317, at 118 (defining external validity as how causal relationships fare when there are changes in persons, settings, or times).

410. "Internal validity" is the extent to which the variable of interest, rather than some other factor, caused the result studied. See id. at 117. See Faigman, Social Science and Policy, supra note 407, at 1060 n.201 ("Sound theory testing in the social sciences invariably involves some amount of laboratory testing which . . . may not be generalizable to the legal setting without further field tests. However, both laboratory tests and field tests have limitations, the former involving external validity (generalizability) and the latter involving internal validity. . . . Yet together, laboratory and field research can provide valuable—and valid—information."). Cook and Campbell in their classic text identify the following seven factors as substantial threats to internal validity:

[History] is a threat when an observed effect might be due to an event which takes place between the pretest and the posttest, when this event is not the treatment of research interest . . . .

[Maturation] is a threat when an observed effect might be due to the respondent's growing older, wiser, stronger, more experienced, and the like between pretest and posttest and when this maturation is not the treatment of research interest.

[Testing] is a threat when an effect might be due to the number of times particular responses are measured. In particular, familiarity with a test can
Second, studies may lack comparison or "control" groups. Ideally, there will be two identical groups, one being subjected to the variable being studied, the other not being so subjected. Thus, any differences between the two groups beyond chance variation can be attributed to the variable manipulated.\textsuperscript{411}

Third, even differences between a subject group and a control group may be due to chance rather than changes in the independent variable.\textsuperscript{412} Statistical formulae must thus be used to determine the statistical significance of any observed differences,\textsuperscript{413} that is, a measure of the likelihood that an observed difference in samples was not due to chance but to the effect of changes in the manipulated variable.\textsuperscript{414}

Fourth, efforts must be made in the experiments to control for "experimenter expectancies," that is, the effect that a researcher's awareness of the hypothesis being tested can have on what and how data is collected.\textsuperscript{415} This problem arises when data contrary to a hypothesis are ignored or when records "reflect what the researcher

---

sometimes enhance performance because items and error responses are more likely to be remembered at later testing sessions.

[Instrumentation] is a threat when an effect might be due to a change in the measuring instrument between pretest and posttest and not to the treatment's differential impact at each time interval . . .

[Statistical regression] is a threat when an effect might be due to respondents' being classified into experimental groups at, say, the pretest on the basis of pretest scores or correlates of pretest scores. When this happens and measures are unreliable, high pretest scorers will score relatively lower at the posttest and low pretest scorers will score higher. It would be wrong to attribute such differential "change" to a treatment because it might be due to statistical regression . . .

[Selection] is a threat when an effect may be due to the difference between the kinds of people in one experimental group as opposed to another . . .

[Mortality] is a threat when an effect may be due to the different kinds of persons who dropped out of a particular treatment group during the course of an experiment. This results in a selection artifact, since the experimental groups are then composed of different kinds of persons at the posttest.


412. Id. at 1061.
413. See id. at 1061-62 ("[Failing to] calculate any measure of statistical significance is a substantial violation of a basic tenet of empirical inquiry.").
414. See MARTIN, supra note 317, at 152 (suggesting that for a result to be significant, the likelihood of obtaining the observed differences due to chance should be less than one in twenty). Most psychologists require a .05 level of significance, meaning that there is a 5% likelihood that any observed difference in samples is due purely to chance rather than to the experimental treatment condition, while others choose a .01 level of significance. See id.
415. See Faigman, Social Science and Policy, supra note 407, at 1062.
expects to hear, rather than what the subjects actually said."

Fifth, what is being measured and its meaning to the researcher must be clearly understood, for sometimes what is studied does not in any real sense test the variable that is of concern to the law.

One important qualification must be noted about these five guidelines. In a true experiment, all variables other than the one to be studied are “controlled,” that is, they are equal for all subjects, or their effects are cancelled out by randomly assigning participants to the various treatment conditions. To some degree, random assignment can assure that the findings result from the factor that the researcher wishes to study, not from pre-existing differences between the groups. True experiments are often impossible in social science research, however, because it is either impossible to control for all variables other than the one being studied or it would be unethical to introduce the independent variable under controlled circumstances. For example, raising twenty children in an identical fashion except that ten are subjected to regular, brutal beatings, would be morally unacceptable.

When true experiments are impossible, social scientists aim to get as close as possible to the ideal by minimizing threats to internal validity or creating situations where those threats can fairly be assessed. A procedure in which there is such partial, but not complete, control over the scheduling of experimental stimuli is called a “quasi-experiment.” Because a quasi-experiment is by definition a flawed experiment, the above guidelines are particularly important in assessing the value of particular quasi-experimental designs.

Finally, when general acceptance of a technique based on scientific experimentation is not the determinative factor for admissibility, as in a pragmatic relevancy jurisdiction, it is nevertheless a factor

416. Id. at 1062-63.
417. See id. at 1063 (noting that in some cases, the measures selected do not reflect the legal concerns or the legal concern is “effectively untestable”).
419. Id.
420. Cf. E. Jerry Phares, Clinical Psychology: Concepts, Methods, and Profession 113 (4th ed. 1992) (“Would it really be ethical to subject such people [depressives] to a strong, significant failure experience and thereby risk plunging them even deeper into depression [just to study the impact of failure on such people]?”).
422. Id.
in weighing probative value.\textsuperscript{423} This makes sense, for scientists recognize that "[t]he trustworthiness of a study increases as independent investigators arrive at a common conclusion. The more often a study is confirmed by subsequent research, the less likely it is that chance fluctuations in the data accounted for the results of the original research."\textsuperscript{424}

This checklist is meant simply to convey the flavor of the enterprise in an individual case and to suggest the ways in which the hallmarks of good social science can easily be applied by courts and practicing attorneys.\textsuperscript{425} Many useful studies will have some of these flaws;\textsuperscript{426} weaknesses detected based upon these and other standards should not, therefore, result automatically in exclusion of the evidence. Rather, the goal is to identify that research which is so inadequate that it will mislead, rather than assist, a jury when weighed against other factors.\textsuperscript{427} Equally important, however, is that even flawed or as-yet-uncorroborated research may be of some evidentiary value, but a jury can accurately gauge that value only if fully informed of both the strengths and weaknesses of the research.

\textbf{ii. Clinical Judgment.—}Evidence that a certain group behaves in a certain way is generally not helpful unless it is further established that the defendant does or does not fit within the group being studied.\textsuperscript{428} Expert opinion offered on this latter question is not, however, subject to experimental or quasi-experimental testing but rather turns on the "clinical judgment" of a relevant mental health professional. Moreover, expert character testimony that is more individualized, emphasizing what is special or unique about a defendant rather than what she shares in common with a group, also requires clinical judgment.

\textsuperscript{423} See McCord, Profiles, supra note 119, at 99 (noting that the greater the acceptance of the accuracy of the diagnosis, the greater chance the diagnosis before the court will be trusted as reliable).

\textsuperscript{424} Monahan & Walker, supra note 418, at 508.

\textsuperscript{425} See id. at 508-12, in which the authors agree that even a "crude screening device," such as "a simple checklist of threats to validity" and of ways to neutralize those threats, "can go far in reducing a court's reliance on inadequate empirical data." Id. at 509. They also emphasize, however, that ease of application will vary with the judge and with the case, and, in some instances, expert consultation will be essential to assist the court. Id. at 509-12.

\textsuperscript{426} See Faigman, Social Science and Policy, supra note 407, at 1065 (noting that most social science research has some flaws).

\textsuperscript{427} See id.

\textsuperscript{428} See supra text accompanying notes 325-330.
"Clinical judgment" is the application of the judgment of a trained mental health professional to an individual case, based on a weighing of data from various sources. That judgment is used primarily for diagnosis, prediction, and designing a treatment plan. The quality of clinical judgment turns on the skill of the clinician in collecting data, for example, in asking the right questions and in promoting responses during interviews, in categorizing and weighing that data, and in drawing reasonable inferences from the data. Moreover, collected data is used to generate hypotheses, which guide the clinician in collecting more data and in revising initial hypotheses until a conclusion is reached. Clinical judgment is thus as much art as science.

The value of clinical judgment has, therefore, often been challenged because it is subject to distorting influences by the clinician and necessarily involves a great deal of scientifically questionable speculation. Although actuarial judgments often prove superior to clinical judgments in predicting behavior, clinical judgment is often all that we have. Moreover, it can offer much of value to the jury by increasing the pool of data and fostering the jury's willingness to consider plausible alternatives. This section seeks, therefore, to serve two purposes: to offer general guidelines for improving—and, in a particular case, for examining—the quality of clinical judgment, and to offer suggestions for improving the accessibility of that judgment to a jury.

429. See Phares, supra note 420, at 262.
431. See Bonnie & Slobogin, supra note 35, at 433-34, 514 (noting that skeptics would likely deny either exculpatory or mitigating significance to claims that psychological forces caused a defendant not to refrain from criminal behavior and that the legal system attempts to avoid bias through its fact-finding and law-applying processes).
432. See Phares, supra note 420, at 267-68. On the other hand, even critics of clinical methodology concede that in the subset of cases for which clinicians express a high degree of confidence in their predictions, clinical predictions may be more accurate than actuarial ones. See John Monahan, Predicting Violent Behavior: An Assessment of Clinical Techniques 120 (1981). Furthermore, clinical prediction may be more accurate than actuarial prediction when dealing with rare factors not taken into account by actuarial formulae. See id. at 119-20. For example, an actuarial formula might predict a 90% likelihood that a professor will attend a movie tonight. But a clinician who knew that the professor had broken her leg would be more accurate in predicting that there was close to a 0% chance of the professor's showing up at that movie. Id. at 119.
433. See infra text accompanying notes 547-563.
434. See infra text accompanying notes 514-534.
In collecting data, it is important that a clinician personally interview the defendant; no courtroom opinion should be offered when this has not been done.\textsuperscript{435} As noted earlier, there should be multiple interviews over several days and in different settings to offset the distorting effect of atmosphere and of a restricted sample of personally observed behavior.\textsuperscript{436} Moreover, the interviews should cover information regarding the defendant’s behavior in a wide range of settings, and some interviews should be done by other clinicians to avoid the distorting effects of a specific interviewer’s personality.\textsuperscript{437} Additionally, the clinician must be aware of the ways in which her own behavior can influence a defendant—“examiner effects”—and must be trained to minimize these effects.\textsuperscript{438} Multiple observers can also help to minimize examiner distortion.\textsuperscript{439}

On the other hand, no opinion should be based solely on what a defendant says.\textsuperscript{440} The accuracy of a defendant’s recounting of her history and thoughts is especially critical because those data will form much of the basis for the expert’s opinion regarding that defendant’s character.\textsuperscript{441} Thus, to the extent possible, the clinician should act as a detective, interviewing other sources of information to confirm, disconfirm, and expand on the information provided by the defendant.\textsuperscript{442} Written data, such as personnel records, should be examined as well.\textsuperscript{443}

Apart from interviewing others, there are further steps that can be taken to test the veracity of a defendant’s self-reports, including:

\textsuperscript{435} Bonnie & Slobogin, supra note 35, at 496.
\textsuperscript{436} Id. at 505.
\textsuperscript{437} Id.
\textsuperscript{438} See id. at 512-14 (suggesting further ways to reduce such effects); Stanley L. Brodsky, Testifying in Court: Guidelines and Maxims for the Expert Witness 79-83 (1991) (defining the term “examiner effects” and suggesting ways to minimize them).
\textsuperscript{439} See Bonnie & Slobogin, supra note 35, at 513 (suggesting the use of one-way mirrors).
\textsuperscript{440} Id. at 508.
\textsuperscript{441} See id. Because this Article focuses on pragmatic relevancy, the hearsay-related issues involved when an expert relies on out-of-court statements made by others, especially those arising under Federal Rule of Evidence 703, are not addressed here. See Fed. R. Evid. 703.
\textsuperscript{442} See Bonnie & Slobogin, supra note 35, at 508-09 (discussing the need to rely on sources other than the defendant and the legal implications of doing so); Ziskin, supra note 239, at 1115 (concluding that a mental health expert must gather and evaluate as much evidence as possible before forming an opinion). Cf. William M. Runyan, Life Histories and Psychobiography: Explorations in Theory and Method 44 (1984) (discussing, in the context of explanatory conjecture, how multiple explanations of behavior may provide more data to reach a “good” explanation).
\textsuperscript{443} Ziskin, supra note 239, at 1105.
(1) using the "validity scales" of the MMPI,\textsuperscript{444} (2) comparing multiple observations of and interviews with a defendant in a search for consistency, and (3) after establishing initial client rapport, adopting—at least in situations where a client has not been as forthcoming as is necessary—a somewhat more confrontational style than is generally used by therapists.\textsuperscript{445} It is important to emphasize, however, that a clinician does not opine about the truthfulness of a defendant's version of the crime. Rather, a clinician only explains why, based upon all her interviews and investigations, she accepted certain data as accurate for purposes of offering an opinion regarding a defendant's character.\textsuperscript{446}

A clinician must also keep detailed records that note observations in concrete terms, not merely as conclusions or generalities, so that a jury can draw its own conclusions from the data.\textsuperscript{447} Indeed, videotaping portions of an investigation, such as interviews, might be an even better way to enable a jury to determine the fairness and accuracy of a clinician's investigation.\textsuperscript{448} Opinions offered by a clinician, moreover, should be framed so that psychological observations and judgments are not confused with legal or moral ones.\textsuperscript{449}

\textsuperscript{444} See supra text accompanying notes 189-199.

\textsuperscript{445} Bonnie & Slobogin, supra note 35, at 505-09. See also Brodsky, supra note 438, at 36-38 (discussing clients' motives to "fake" their own evaluations). Other potential "truth-testing" techniques, such as sodium amytal, hypnosis, and the polygraph, raise complex evidentiary questions in their own right, see id. at 506, but are beyond the scope of this Article.

\textsuperscript{446} For a discussion of the special problems created when an expert offers an opinion about the truthfulness of a witness or an accused, see McCord, Profiles, supra note 119, at 44-48.

\textsuperscript{447} See Ziskin, supra note 239, at 1104.

\textsuperscript{448} See id. at 1105-06. This last suggestion raises problems of its own, however, such as whether the videotaping may itself alter a defendant's behavior, the potential need for defense consent to such a procedure, id., and Fifth Amendment concerns in the absence of such consent. See Bonnie & Slobogin, supra note 35, at 497-503, 507-08 n.238 (noting that the most powerful disincentive to full disclosure is the fear that such information may be used against the defendant and the related notion that the use of a videotape as evidence may run afoul of a defendant's Fifth Amendment rights). See also Shuman, supra note 189, at 282-84, 1990 Supp. at 52-54 (discussing the Fifth Amendment's limitation on a psychiatrist's opinion generally and the use of videotaping specifically). Nevertheless, there is significant support for videotaping, see, e.g., Runyan, supra note 442, at 157 (discussing the usefulness of videotaping interviews), such that it should be available as a potential aid when a defendant consents. See Shuman, supra note 189, at 238.

\textsuperscript{449} David L. Bazelon, Questioning Authority: Justice and Criminal Law 59 (1987) ("In practice . . . psychiatrists have continued to make moral and legal judgments beyond the proper scope of their professional expertise."). See also Smith v. Schlesinger, 513 F.2d 460, 462 (D.C. Cir. 1975) ("[P]sychiatric judgments may disguise, wittingly or unwittingly, political or social biases of the psychiatrists.").
Wherever possible, clinicians also should make use of whatever actuarial formulae are available, should be sensitive to relevant base rates, and should be sensitive to the apparent effects of differing situations on a defendant's behavior. Ideally, a prediction should indicate in what situations, for example, a defendant is most likely to be violent and why. The clinician must further be sensitive to, and look for, data that may disconfirm her working hypotheses. Such attention allows the clinician to make judgments in part based on an "objective" comparison of the defendant to others, while enabling her to defend those judgments by recognizing and rebutting factors that may undermine confidence in them. As a matter of practice, the level of confidence a jury should place on the opinion should be stated as part of the opinion; for example, "it is more likely than not that the defendant will be violent if returned to situation X."

Additionally, the data collection and interpretation process should be an ongoing one, constantly subject to review and to critique by other clinicians. This means that all hypotheses should be tentative and that a group approach is preferred over the use of a single clinician. Multiple viewpoints make it more likely that weaknesses will be uncovered, missing data will be collected, and coherencies will be found. Videotape, again, may help to permit multiple interpretations of the same data.

Furthermore, concerns may be raised because of "fallacious prediction principles"—failures to recognize the application of simple rules for good statistical reasoning—made by laypersons and clinicians alike. One example would be the tendency to judge a

450. See Monahan, supra note 432, at 49, 121-27, 129-41, 160. Monahan attacks the particularly thorny problem of predicting dangerous behavior, an area in which he concedes psychologists have an abysmal record. See id. at 27-28. However, he argues that the poor record has resulted partly from inadequate procedures and methods and suggests appropriate methods for this challenging task. See id. at 19, 37, 143-69.

451. See id. at 157-58 (describing a prediction model the clinician should use to determine the probability of violent behavior).

452. See id. at 160 (emphasizing the need to corroborate evidence).

453. See, e.g., id. at 34, 168.

454. See Bonnie & Slobogin, supra note 35, at 512-22 (describing a detailed team approach to quality control for clinical assessments, based upon the model used at the University of Virginia's Forensic Psychiatry Clinic); Runyan, supra note 442, at 157 (suggesting that multiple viewpoints of various clinicians can serve as a way of checking and rechecking initial clinical descriptions and interpretations).

455. See Bonnie & Slobogin, supra note 35, at 516-17 (noting that the segments of a team interview provide additional viewpoints on the data, increasing the likelihood that all facts and theories will be considered).

person by a single instance of extreme behavior rather than by a broader and fairer sample of behavior.\textsuperscript{457} A thorough clinician must be aware of such errors, carefully guard against them, and stand ready to explain to a jury the method by which the investigation was conducted and conclusions were drawn.

These guidelines are offered to help courts determine when a particular exercise of clinical judgment—particular psychological character evidence—is so defective as to require exclusion. When such expert testimony is admitted, these guidelines may aid in ensuring full disclosure to the jury of the significant strengths and weaknesses of the character evidence.

c. \textit{Witness Qualifications}.—Forensic psychiatry and psychology are best viewed as specialties requiring appropriate education and training. There are a growing number of routes available for formal classroom and clinical training in this area,\textsuperscript{458} but forensics-specialist programs are often not very rigorous or complete.\textsuperscript{459} While awaiting improvements in the formal university training programs, courts must be sensitive to the special type of forensic training that is necessary to testify as a specialist.

Absent such training, many clinicians are simply not fit for the courtroom. Commentators suggest that “their training encourages them to err in the direction of diagnosing illness, . . . to speculate wildly about unconscious determinants of behavior,” and to ignore the limitations of their own disciplines.\textsuperscript{460} To be able to provide juries with useful and relevant information, a forensics expert must have some basic knowledge of the relevant substantive and procedural law, including the rules of evidence, enabling them to appreciate the differences between psychological and legal reasoning. The forensics expert ought to be skilled in careful, thorough, concrete, and precise record-keeping that offers particularized support for

\textsuperscript{457} See \textit{Phares}, supra note 420, at 270, 272 (describing such clinician prediction fallacies).

\textsuperscript{458} See \textit{id.} at 484 (discussing the availability of joint degree programs in law and psychology, Ph.D. programs offering specialization in law and psychology, and university and forensic mental health programs offering postdoctoral training opportunities).

\textsuperscript{459} While a 1973 survey of 83 university medical centers indicated that 90% of these schools had some type of combined law and psychiatry program, “very few postgraduate or post-residency fellows were being trained, . . . most courses used informal notes and case materials rather than a standard text, and . . . little field work was required.” Bonnie & Slobogin, supra note 35, at 457-58 n.103.

\textsuperscript{460} \textit{Id.} at 457.
opinions they will offer in court. Moreover, because an expert’s task is not to cure but to evaluate in terms of legal standards, he must be familiar with the different procedures necessary for conducting a quality forensic evaluation, most notably the need to engage in extensive inquiries of third parties to confirm, disconfirm, or supplement a defendant’s tale. Experts also must be trained in how to respond appropriately to cross-examination and, while being persuasive, must learn not to be more than a “mild advocate,” able to defend a position without letting the lawyer’s need to win affect their professional judgment. Significant experience in a clinical setting will obviously enhance an expert’s qualifications. Of course, individual jurisdictions will require at least minimal credentials as well, generally either board-eligibility as a psychiatrist or a Ph.D. in clinical psychology.

Beyond the requirements for recognition as a forensics expert, an expert must also have the necessary theoretical and experiential training in the particular subject on which he will testify. Whether the subject is the battered woman syndrome, child sexual abuse accommodation syndrome, or a more individualized inquiry based on the MMPI and theories of individual schemata, an expert must be thoroughly versed in the relevant literature, aware of the strengths and weaknesses of the data, and practiced in applying that data to real-world situations. Finally, an expert must demonstrate knowledge of the necessary steps in a thorough evaluation procedure in the particular area and show an ability to explain why, in that case, he did or did not engage in any of the steps that might be part

461. These suggested qualities for forensics experts are adapted from a well-known psychiatrist’s list of qualities for the courtroom-wise witness. See Brodsky, supra note 438, at 134-35. Others have suggested that forensic training must ensure a “familiarity with legal tests and concepts, proper assessment, knowledge of relevant literature, and an orientation to the courtroom.” Phares, supra note 420, at 484. Minimal training should, therefore, include an introductory survey course in forensic psychology, topical seminars in the area, and a field placement in a forensic setting. Id.

462. Brodsky, supra note 438, at 135.


464. Id. at 457 n.103. The courts’ greater willingness to qualify psychiatrists over psychologists for issues related to the mental state of a defendant is unjustified unless somatic conditions are at issue. See id. at 458-59 n.105. Professors Bonnie and Slobogin also suggest that a psychiatric social worker with a master’s degree in social work or a master’s level psychologist or psychiatric nurse “will ordinarily have sufficient clinical-behavioral training to be helpful in some legal contexts, such as competency assessments, so long as other qualifications are established.” See id. at 457 n.103.

465. See, e.g., Monahan, supra note 432, at 146 (noting several ways to determine if an expert has the necessary training to predict the probability of future violence).
of an "ideal" examination. An expert's aptitude for conveying such an explanation is a strong indication of whether he will be able to fully inform a jury of the facts and relevant arguments that should affect its deliberations.

d. Negative Inferences.—Expert character testimony is often offered to support a "negative inference"—an inference derived from proving the absence, rather than the presence, of a certain trait or quality. This proof of a negative may be challenged as having little or no probative value.

Even a brief foray into daily newspaper stories, however, demonstrates the weakness of this challenge. The much-publicized allegations of Anita Hill at the confirmation hearings of Justice Clarence Thomas provide a good example.

Despite the presence of corroborating witnesses, many were simply unwilling to believe Professor Hill's story. But if her story was untrue, the only explanation was that she was either lying or emotionally disturbed. Some inferred that she "must" have been a "spurned woman," eager for revenge against a man who slighted

466. See Bonnie & Slobogin, supra note 35, at 460 ("Because the reliability and usefulness of an otherwise qualified expert's testimony depends largely on the quality of his evaluation procedure, courts also should pay special attention to this subject. Ideally, courts should make this assessment before qualifying the witness . . . .").

467. Professor Saltzburg has defined a "negative inference" as an inference "drawn by factfinders from the absence of one or more pieces of evidence that they expect to be presented." Stephen A. Saltzburg, A Special Aspect of Relevance: Countering Negative Inferences Associated with the Absence of Evidence, 66 CAL. L. REV. 1011, 1011 (1978). While Professor Saltzburg's definition and examples differ slightly from those offered here, both Professor Saltzburg and I share a common concern: that jurors may draw inappropriate inferences based on assumptions that can be countered, if at all, only by admitting evidence specifically intended to prove those assumptions incorrect. See id. at 1060.

468. Anita Hill, a well-spoken, intelligent law professor at the University of Oklahoma testified that while she was supervised by Clarence Thomas at the Equal Employment Opportunity Commission, he engaged in a series of bawdy conversations with her that she later characterized as sexual harassment. See CLARENCE THOMAS, CLARENCE THOMAS: CONFRONTING THE FUTURE 59 (L. Gordon Crovitz ed., 1992) (describing specific allegations Hill levied against Thomas); Ruth Marcus, Hill Describes Details of Alleged Harassment; Thomas Categorically Denies All Her Charges; Court Nominee Calls Ordeal "Lynching for Uppity Blacks," WASH. POST, Oct. 12, 1991, at A1 (describing Hill's allegations of sexual harassment).


470. The arguments recounted here about the respective credibility of now-Justice Thomas and Professor Hill, including the arguments about what "must have" motivated Professor Hill, were made by students, faculty, and friends to this author during the course of the hearings. I suspect that these discussions were typical of those taking place throughout the nation. See also sources cited infra notes 471-473.
her. These inferences, made even though no evidence supported them, could have been countered, at least to some degree, by evidence that the assumed facts—being "spurned" or being "emotionally disturbed"—did not exist. In particular, a psychologist might have testified that Professor Hill displayed neither emotional instability nor unusual personality traits to suggest that she would lie or fantasize about the events to which she testified.

If such evidence had been offered, the outcome would not necessarily have been different. Nevertheless, evidence of the absence of emotional instability, if believed, would have made further speculation difficult and might have changed a few minds. If a similar dispute arose in the context of a criminal trial, however, changing a few minds could make the difference between conviction and acquittal.

In the context of lay character testimony, one recent case recognized the importance of negative inferences. In Barlow v. Commonwealth, the defendant was convicted of murder for fatally shooting his father. Barlow claimed self-defense and argued on appeal that the trial court erred by excluding lay testimony that he "ha[d] no reputation for violent behavior in the community." The state argued that the absence of a reputation for violence is not the same as affirmative testimony of a good reputation for peacefulness, and that

471. "Did she [Professor Hill] fantasize about Judge Thomas and then was furious because she felt deserted by him? Is she a militant feminist determined to sacrifice truth in the overall fight against sexual harassment? No evidence has surfaced to indicate such motives." As the Tragedy Continues, HARTFORD COURANT, Oct. 13, 1991, at D2 (emphasis added).

472. Several witnesses in Justice Thomas's favor did speculate as to the existence of some of the assumed facts, but the press noted that such speculation was little proof, even in a Senate hearing. See Adam Clymer, The Thomas Nomination; Parade of Witnesses Support Hill's Story, Thomas's Integrity, N.Y. TIMES, Oct. 14, 1991, at A11 (noting that Phyllis Berry, a witness for Justice Thomas, expressed the view that Anita Hill had a crush on Thomas, but was unable to point to any specific reasons for believing this). These "facts" remained assumptions and nothing more.

473. The closest thing to psychological testimony in the hearings was Professor Hill's report that she had passed a lie detector test, although it was also reported that Republicans contacted Park Dietz, a psychiatrist, to question him about delusions and false allegations of sexual misconduct. See Clymer, supra note 472 (discussing Professor Hill's lawyers' comments that the polygraph illustrated Hill's truthfulness about Justice Thomas's conduct toward her and the Republicans' contact with Dr. Dietz).

474. It is important to remember that the fact-finder in the Thomas-Hill dispute was a United States Senate committee and not a traditional jury. The hearings, as discussed here, are meant to be merely illustrative of the probative value of psychological testimony and negative inferences.

475. 297 S.E.2d 645 (Va. 1982).

476. Id. at 645-46.
only the latter should be admissible. Moreover, argued the state, character evidence was irrelevant because Barlow had admitted shooting his father.

The Supreme Court of Virginia rejected both prosecution arguments and ordered a new trial. The court described the distinction between the absence of a reputation for violence and the presence of a reputation for peacefulness as merely semantic. The court also emphasized that, although there was no dispute that the son killed his father, the character evidence might negate the necessary mental state for the crime.

The Barlow court thus recognized that a jury might be aided by evidence of a defendant's lack of a reputation for violence, if indeed he had not displayed violent tendencies in the past. Without this exculpatory or mitigatory evidence, a jury might conclude that the defendant "must" have a violent character, and therefore the necessary mental state, because "normal" sons do not shoot their fathers. Absent evidence to the contrary, no other explanation may arise to contradict the jury's usual assumptions. Moreover, jurors might even be more reluctant to believe the son's self-defense claim because, for some, it is hard to believe that a father would seriously threaten his son. Without character evidence illustrating the absence of a reputation for violence, the jury might draw inferences not well-supported by the evidence, and conjure up its own facts to "fill in the blanks" and create a believable "story."

An expert character witness might, for similar reasons, offer testimony of even greater probative value than that offered by lay witnesses. How much probative value, of course, depends upon the details of the testimony. If it were possible, for example, for an expert to testify that, with rare exceptions, only persons fitting a "rapist profile" commit rapes, and the defendant did not fit such a profile, such testimony should be of significant probative value. But the more common situation will be one in which such unequivocal testimony is unavailable. Nevertheless, more equivocal testi-

477. Id. at 646.
478. Id. at 646-47.
479. Id. at 647.
480. Id. at 646.
481. Id. at 647.
482. See infra notes 535-546 and accompanying text (discussing role of stories in jury deliberations).
483. See State v. Cavallo, 443 A.2d 1020, 1023, 1030 (N.J. 1982) (noting that psychiatrist's testimony that a defendant lacked the psychological traits of a rapist was logically relevant, although the testimony was properly excluded on other grounds).
mony—for example, that there is nothing abnormal in the defendant’s personality profile—may be relevant for a different reason: without testimony that jurors expect or want to hear, they will draw inferences that should not fairly be drawn from the evidence, or they may give fair inferences undue weight.

For example, an adult male who enjoys babysitting and spending time with children might be accused of sexual abuse. Because many people may find it unusual for an adult to enjoy spending so much time with children, jurors might speculate that he is a sexual psychopath or a pervert, making it more likely that they will convict. But expert testimony explaining that the defendant displays no signs of unusual behavior or unusual thinking concerning sex generally—and children specifically—may help to counter such common inferences. If the testimony further includes a healthy, relatively “normal” explanation based on the defendant’s background for why he enjoys babysitting, that may go even further to legitimately raise a reasonable doubt. Such testimony will be relevant even if many or most people who commit child sexual abuse are indeed not perverts or psychopaths, because it helps to counter inferences that the jury has little right to draw.

2. Necessity.—“Necessity” directs courts to consider the need for proffered testimony in the context of a specific trial. The need for expert psychological character testimony turns on an understanding of the political role of the jury, the role of suppositional science and storytelling in leading jurors to consider plausible alternatives to normal, the value of the clinical arts as opposed to science, and the absence of less objectionable alternatives.

a. The Political Role of the Jury.—The jury is an inherently political institution, in that it serves important values connected with the role and legitimacy of government in the American democratic sys-

484. Cf. People v. Stoll, 783 P.2d 698, 715 (Cal. 1989) (holding that psychologist’s opinion that defendants displayed no signs of “deviance” or “abnormality” should have been admitted as relevant to defense claim that charged acts of child abuse did not occur).

485. McCord, Profiles, supra note 119, at 95.

486. The last of these factors is expressly articulated by Professor McCord, and several of the other factors follow logically from his explanation of “necessity.” See id. at 95-98. The remaining factors were not suggested by Professor McCord because they follow from an understanding of the cathartic function and the moral and social psychological bases for the law of criminal evidence, matters not considered in his article.
There are at least three primary political values connected with the jury's role that can be promoted by carefully crafting rules of evidence: (1) operational accuracy, (2) the independence of the trier of fact from governmental control, and (3) fairness between the parties. The need to serve these political values should be an integral part of the "necessity" determination under a relevancy approach.

The first of these values—"operational accuracy"—is rooted in the following assumption: rarely, if ever, can we verify in some objective fashion what the "truth" is in a specific case and, therefore, whether we have achieved an "accurate" result. Because we cannot determine whether a judgment is indeed accurate, the best we can do is maximize the likelihood of an accurate result by improving the trier's ability to apply its own general knowledge and experience to an individual case. To do so, the fact-finder must have broad access to relevant information. Operational accuracy can be achieved, therefore, by adopting procedures and evidentiary rules that will enhance the trier's function by increasing the quantity and quality of "foundation facts"—those facts that better enable the trier to judge the trustworthiness of other evidence.

Professor Swift advocates achieving operational accuracy through modification of the hearsay rules. Instead of general rules that prejudge the accuracy of certain classes of hearsay evidence in most cases, she argues for rules designed to promote the availability to the jury of adequate information regarding the foundation facts. This would enable the jury to determine a hearsay declarant's opportunity to observe, remember, tell the truth, and communicate the testimony effectively.

---

487. See Alexis de Tocqueville, Democracy in America 127 (Richard D. Heffner ed., Mentor Paperback 1956) (1835) ("The jury is, above all, a political institution, and it must be regarded in this light in order to be duly appreciated."); James P. Levine, Juries and Politics 20 (1992) ("Trials are more than fact-finding exercises; they entail conflicts of values, interests, and ideologies. The jury, from this perspective, is an inherently political institution.").

488. See Eleanor Swift, A Foundation Fact Approach to Hearsay, 75 Cal. L. Rev. 1339, 1342-43 (1987). Professor Swift, who first identified the notion of "operational accuracy," describes it as a value related to the "political role of the factfinder," id. at 1367, but does not expressly describe all the remaining values as inherently "political" ones; that is nevertheless what they are. See supra note 487 and accompanying text (defining "political").

489. See Swift, supra note 488, at 1342.

490. See id. at 1362-63.

491. See id. at 1355-56, 1361-67.

492. Id. at 1361-67.
Professor Swift wrote of the value of promoting operational accuracy in the fact-finder generally.\textsuperscript{493} That value, however, is even greater when the fact-finder is a jury rather than a judge. Human experience, not legal training, is the most necessary training in making the credibility judgments required in trials.\textsuperscript{494} Moreover, a group of jurors is more likely collectively to remember the details of testimony, and a pooling of opinions in a collaborative decision-making process provides a way to test ideas and distinguish sense from nonsense.\textsuperscript{495} Perhaps most importantly, the lack of an objective truth and the often-vague legal standards applicable to the facts inevitably require a fact-finder to bring values to bear in resolving complex questions of fact.\textsuperscript{496} Jurors as a group offer a wider wealth of experience than an individual judge, and it is the jury who best represents community values.\textsuperscript{497} This is especially important in criminal cases where the \textit{mens rea} requirement ultimately impels jurors to determine moral culpability.\textsuperscript{498} Thus, in criminal cases the value of "operational accuracy" deserves special weight.

A second political value promoted through the rules of evidence is the independence of the trier of fact from government control. General governmental rules applicable to trials are based on a host of considerations, many of them bureaucratic ones, that are unconnected with an individualized assessment of the case before the fact-finder.\textsuperscript{499} The independence of the jury from this government control promotes more individualized justice and is an important political value engendered in the role of the jury.

\textsuperscript{493} Id. at 1361-62.

\textsuperscript{494} See Levine, supra note 487, at 182. Professor Swift makes a similar point but is more willing to trust the judge's as well as the jury's "distillation of experience." Swift, supra note 488, at 1362-63.

\textsuperscript{495} Levine, supra note 487, at 182.

\textsuperscript{496} See id. at 11-17 (noting sources of jury discretion); see also Harry Kalven, Jr. & Hans Zeisel, \textit{The American Jury} 165 (1966) (describing the "liberation hypothesis": "The closeness of the evidence makes it possible for the jury to respond to sentiment by liberating it from the discipline of the evidence").

\textsuperscript{497} See Levine, supra note 487, at 17 ("[T]he jury to a considerable extent mirrors popular values and sentiments . . . . [I]t is a means for the expression of majoritarian sentiment."). See also Ballew v. Georgia, 435 U.S. 223 (1978), in which Justice Blackmun concisely expressed this idea in addressing whether the size of criminal juries may be smaller than six members:

Because juries frequently face complex problems laden with value choices, the benefits [of the larger group] are important and should be retained. In particular the counterbalancing of various biases is critical to the accurate application of the common sense of the community to the facts of any given case.

\textit{Id.} at 233-34.

\textsuperscript{498} Levine, supra note 487, at 12, 82.

\textsuperscript{499} Swift, supra note 488, at 1363.
Independence of the jury is furthered by the concept of nonaccountable decision making and the use of general, nonprecedential knowledge. Nonaccountability furthers independence because it permits the fact-finder to function efficiently without accounting for how its inferences were made, thus preventing judicial review of the generalizations applied by the jury. Such review would engender disagreement about those generalizations and create a risk that judges would substitute their own generalizations for the jury's. Uncertainties about judicial second-guessing would also undermine the willingness of juries to engage in independent behavior.

"Nonprecedential generalizations" are based on general knowledge from jurors' experience, rather than on prior similar cases. They allow for jury independence and promote individualized justice in ways that prior, judge-imposed generalizations cannot. This is because the latter must give great weight to government interests or other established policies beyond the specific case. Because juries take into account only their own limited knowledge of the wider impact of legal rules, they are more likely than a judge to focus only on what is before them—to individualize the trial. As discussed throughout this Article, however, there are strong forces at work in criminal cases—specifically, limited resources, stereotypes, and social myths—that promote bureaucratic decision making by both judges and juries. Nonetheless, the logic and underlying moral assumptions of the substantive criminal law demand individualized justice. To the extent expert character evidence aids in protecting a jury's independence, the two forces of nonaccountable decision making and nonprecedential generalization will act together to promote individualized justice. This is a clear justification for admitting expert psychological character evidence.

500. Id. at 1367. Nonaccountability is the idea that the jury is not required to account for, and is generally not negatively affected by, its reliance on particular sources of knowledge or its evaluation of credibility. Id.
501. Id. at 1367-68.
502. Id. at 1368.
503. Id.
504. Id.
505. Id.
506. Professor Swift argues, to the contrary, that it makes "no difference" to this analysis whether the judge or jury is the fact-finder, for a judge in a bench trial is expected to act as a lay fact-finder. Id. at 1369. Relative to a jury, however, it is unlikely that a judge can ignore the impact of her decisions on the wider legal system—even if the resulting effect on her decisions is a subconscious one—for systemic considerations are part of her daily training and experience.
507. See supra notes 56-118 and accompanying text.
The third value—achieving fairness through trial-access to proof—is important for several reasons. Fairness between the parties legitimizes trial outcomes in an adversary system.\(^{508}\) Additionally, "[t]he proponents' fairness claim is to be allowed to perform their trial function—proving their own cases and enforcing the substantive law, with the freedom to use the widest range of accessible evidence. The opponents' fairness claim is to have access to the proponent's sources of proof at trial."\(^{509}\) These fairness concerns are greatest in criminal trials, where so much is at stake and where moral opprobrium attaches to a conviction. In the specific case of expert character evidence, encouraging the fullest possible presentation of evidence by the proponent, and the fullest availability to the opponent of the proponent's sources of proof, is especially critical. The parties have a need to leave the trial feeling that they had a real opportunity to present an individualized picture of what happened, and to be judged by a jury fully informed of the defendant's character and particular circumstances.\(^{510}\) This cathartic function\(^{511}\) is thus best served by increasing the availability of character information to the jury.

An approach to crafting relevant evidentiary rules that requires the proponent to come forth with as much information as is reasonably possible regarding the value of the psychological character evidence to be presented is, therefore, preferred over one that errs on the side of categorically excluding much of the expert character evidence. This approach also suggests that fair disclosure between the parties, preferably pretrial,\(^{512}\) is required if the exploration of the value of psychological character evidence at trial is to be meaningful. This will also serve the related notion that adequate disclosure in and of itself is an aspect of fairness.\(^{513}\)

In short, each of these political values suggests that the evidentiary analysis used to determine the admissibility of expert psychological character evidence should promote the availability of such evidence, both in quantity and quality. In addition, information relevant to estimating the value of the testimony must be available to a jury. The need to serve these political values must be given special weight in the balancing of probative value against countervailing

\(^{508}\) See Swift, supra note 488, at 1969.  
\(^{509}\) Id. at 1969-70.  
\(^{510}\) See supra notes 56-118 and accompanying text.  
\(^{511}\) See supra text accompanying notes 331-345 (defining the "cathartic function").  
\(^{512}\) See infra text accompanying notes 642-648.  
\(^{513}\) Swift, supra note 488, at 1374.
considerations required by a relevancy analysis of expert character evidence.

b. Suppositional Science, Plausible Alternatives, and Storytelling.—

i. Suppositional Science as a Way of Suggesting Plausible Alternatives to Normal.—Professor David Faigman argues that "suppositional science"—scientific theory based on untested or inadequately tested empirical data—should never be admitted for consideration by a jury. He also argues, conversely, that suppositional science should often be used by legislators and judges in crafting legal rules. He points out that judges often use data that is not falsifiable, such as historical data, in creating law; to interpret the data, however, courts can choose among disputed theories based upon which one best fits all the available data and thus has more persuasive force. Professor Faigman further argues that when theories are subject to legitimate dispute, lawmakers should give the underlying hypotheses due consideration because they can provide helpful insights that lawmakers might otherwise miss, and because they are at least as valid as a lawmaker's best guess. While values may also come into play in choosing among competing but inadequately proven theories, Professor Faigman believes this is appropriate, because making choices among competing values is part of a lawmaker's job. Finally, while there is a danger that lawmakers will misunderstand and thus misuse suppositional science, Professor Faigman argues that any such prejudice can be cured by educating them about its strengths, weaknesses, and limitations.

514. Faigman, Social Science and Policy, supra note 407, at 1052. Professor Faigman states, specifically, that suppositional science refers to findings that "on their face are untestable or have not been tested in any fashion whatsoever; and . . . those that assume the veneer of science (i.e., are forwarded as fully tested propositions) but have yet to be tested adequately." This is a description that fits much clinical psychological testimony. See Bonnie & Slobogin, supra note 35, at 430 (acknowledging the particular imprecision of assessing the "range of variables influencing human behavior").

515. Faigman, Social Science and Policy, supra note 407, at 1078-79 ("[L]awmakers should have the benefit of a variety of views, both objectively scientific as well as subjectively persistent. The juror's role, in contrast, is limited strictly to that of applying the law to the facts as they are determined to exist.").

516. Id. at 1069; see also infra notes 547-574 and accompanying text.

517. Faigman, Social Science and Policy, supra note 407, at 1070.

518. Thus Professor Faigman warns that while the suppositional scientist's own value judgments are entitled to "respect," the ultimate weight of those values is for the lawmaker, who must make an "independent legal judgment of the force of these values." Id. at 1070-71.

519. Id. at 1071.
Rephrased, Faigman suggests that suppositional science has value to lawmakers because it suggests plausible alternatives that might not otherwise be considered. Properly presented, it also educates them effectively in a way that better enables them to bring their values to bear on an uncertain world, to make the best choice among competing alternatives. But expert character evidence, some of which is indeed "suppositional science," can play a similar role for the jury. Such evidence suggests plausible alternative interpretations to the "normal" ones that are likely to occur to the jury, and provides jurors with information that better enables them to choose among competing alternatives. It is true that jurors' values will come into play, but a true understanding of the role of the jury must acknowledge that when jurors have any discretion—and discretion is greatest in cases with conflicting evidence—their values will inevitably affect their decisions. Judges sitting as fact-finders will likewise bring their values to bear in interpreting ambiguous facts, yet a jury's verdict is more likely to reflect fairly the community's values rather than those of an individual judge. Admittedly, some values that we do not deem appropriate for juries to consider—values based on prejudices, stereotypes, and myths—may lead to the misuse of expert character evidence, but as with lawmakers, jurors can be effectively educated to reduce the likelihood of such misuse. Indeed, one major purpose of expert character evidence in the first place is to prevent juror misuse of discredited "quasi-evidence," using educational character evidence to overcome pre-existing biases that endanger individualized justice.

Professor Faigman also unfortunately casts his net too widely—seeking to apply his approach to civil and criminal cases alike—and ignores the special functions unrelated to truth-seeking that rules of evidence can sometimes serve, especially in the criminal context. As noted throughout this Article, expert character evidence can serve a "cathartic function," rooted in the moral dictates of the criminal law and in moral psychology, and can serve certain political values other

520. See supra text accompanying notes 485-513; infra text accompanying notes 535-565.
521. See Bonnie & Slobogin, supra note 35, at 485 (noting that clinical testimony on mens rea provides a plausible alternative to "normal" assumptions about conscious motivation).
522. See supra text accompanying notes 487-513.
523. See supra text accompanying notes 493-498.
524. See infra text accompanying notes 616-647.
525. See supra text accompanying notes 11-94.
Moreover, even when expert character evidence is based on "suppositional science," as long as the alternatives suggested by the evidence are plausible, truth-seeking is promoted when the testimony helps to overcome preconceptions that might otherwise prevent a fair consideration of all plausible interpretations of the evidence. When there is a strong need for evidence and little danger of "overawing" the jury—and when fears of prejudice can be adequately addressed by jury education—there is little justification for flatly excluding even suppositional science from the jury's decision-making process.

Professor Faigman considers most clinical decision making to be suppositional science. Much expert character evidence, it should be recalled, involves classifying a defendant as a member of a group whose members are more likely than nonmembers to engage in certain thoughts or behaviors. The classification process is accomplished by the kind of clinical decision making that Professor Faigman abhors. While Professor Faigman would admit valid—that is, nonsuppositional—data regarding group behavior, he would generally exclude the mental health professional's opinion that a defendant is indeed a member of that group. He would also allow the clinician to testify about actual observations of a defendant's behavior. Professor Faigman asserts, however, that no evidence supports the idea that clinicians are any better than laypeople at classifying particular individuals, so we should let the jury decide for itself.

Assuming that Professor Faigman is correct about the absence of proof regarding the special abilities of clinicians, his argument is flawed for several reasons. First, studies strongly suggest that abstract, general testimony that is not linked to the defendant will

---

527. See supra text accompanying notes 293-309.
528. See Faigman, Social Science and Policy, supra note 407, at 1072-78.
529. Id. at 1076.
530. In so allowing, Professor Faigman is apparently accepting Professor Morse's argument that clinicians are better observers than lay people because relevant data collection of this sort is what clinicians do every day. Id. at 1076 n.265; see Richard Morse, Crazy Behavior, Morals, and Science: An Analysis of Mental Health Law, 51 S. Cal. L. Rev. 527, 601, 611 (1978) (acknowledging the role of hard scientific data and observations of behavior in the clinician's assessment of whether a defendant is crazy, but arguing that such a clinical diagnosis fails to convey accurate and specific information about the defendant's behavior).
531. See Faigman, Social Science and Policy, supra note 407, at 1076 ("'There is no reason to believe that triers of fact, when provided with relevant studies, cannot assess the particular defendant's dangerousness as well as an expert.'").
barely be noticed by the jury; therefore, for evidence to have a significant impact on jury deliberation, testimony about group behavior must be linked to the specific case.\(^{532}\) There is no reason to exclude this if, as clearly may be supposed, clinicians are no worse at this task than are laypeople. Second, Professor Morse, upon whom Professor Faigman in part relies, concedes that descriptions of “reasoning and control processes,” which apparently include such character traits as “impulsiveness” and “suspiciousness,”\(^{533}\) can be among the “observations” that mental health professionals are better able to make than laypersons. Third, mental health professionals and their theories can help to organize data in ways that offer jurors new insights that they might not draw on their own. This last point can best be understood by exploring the conception of how jurors reason that is embodied in “storytelling theory.”

ii. Storytelling Theory.—“Storytelling theory,” which is supported by a significant body of empirical data,\(^{535}\) holds that juries deliberate by constructing stories based on the evidence presented to them at trial.\(^{536}\) Walter Fisher’s theory of narrative rationality is the most useful version of storytelling theory for present purposes.\(^{537}\) Drawing from Fisher, Professors Rieke and Stutman maintain that the story of a case must be told in such a way as to satisfy a jury’s need for “narrative coherence” and “narrative fidelity.”\(^{538}\)

“Narrative coherence” concerns whether the story “hangs together,” a question that is assessed by whether the story has argumentative or structural coherence; material coherence, which involves comparing and contrasting with other stories with which the jurors are familiar, in a search for problems, counter-arguments, and missed issues; and characterological coherence, which addresses “the central feature of all stories, character.”\(^{539}\) A jury is

---

532. See sources cited supra note 328.
533. Morse, supra note 530, at 616.
534. Bonnie & Slobogin, supra note 35, at 455 (interpreting Professor Morse’s article).
535. See SPECIAL COMMITTEE ON JURY COMPREHENSION, ABA SECTION OF LITIGATION, JURY COMPREHENSION IN COMPLEX CASES app. 10, at 18-19 (Dec. 1989) (hereinafter ABA, JURY COMPREHENSION) (citing authorities).
536. See id. For an interesting explanation of how storytelling shapes our memories and is intricately tied up with how we reason, see generally Roger C. Schank, Tell Me A Story: A New Look at Real and Artificial Memory (1990).
538. Richard D. Rieke & Randall K. Stutman, Communication in Legal Advocacy 94-95 (1990). The authors also effectively summarize some of the leading alternative storytelling theories. See id. at 98-102.
539. Id. at 94.
most likely to doubt character if a person’s actional tendencies are contradictory or change significantly from one point of the story to another.\textsuperscript{540}

“Narrative fidelity” is demonstrated through the logic of “good reasons.”\textsuperscript{541} Good reasons are the means by which jurors decide if the story of a case is accurate, in terms of conformity to the jurors’ sense of reality. Good reasons fuse logic and values. Thus, narrative fidelity may involve: (1) whether relevant facts have been omitted or distorted, and the nature of the values implied by the facts in a case, (2) whether the connections of facts to conclusions are reasonable, and whether jurors can accept that purported rationality without violating their own sense of self, (3) whether the reasoning implied in connecting facts to conclusions is all that should be considered in a case, and whether the values implied are confirmed by the jurors’ own experience, and (4) whether the story addresses the “real” issues in the case, and whether the values involved constitute the ideal basis for human conduct.\textsuperscript{542}

An example offered by Professors Bonnie and Slobogin in a different context illustrates the application of these principles to expert character evidence.\textsuperscript{543} Ms. B was charged with knowingly possessing nine stolen welfare checks. The prosecution’s undisputed evidence showed that six checks, made out to different payees, were given to Ms. B by her boyfriend’s good friend, Scott. Ms. B deposited the checks, which supposedly had been endorsed to Scott, into her account. One check was returned unpaid and her account charged accordingly, but Scott made good her loss. He subsequently deposited three more checks. The prosecution’s theory is that Ms. B “must have known” that the checks were stolen, but Ms. B testified that she did not in fact believe that the checks were stolen; she believed Scott’s story that the checks had been given to him in payment for a debt, and he needed her assistance because he had no bank account of his own.

Standing alone, this evidence would probably lead a jury to reject Ms. B’s story and convict her as charged. Her story lacks narrative coherence because it is inconsistent with jurors’ preexisting

\textsuperscript{540} Id.
\textsuperscript{541} See Fisher, supra note 537, at 105-10; see also Rieke & Stutman, supra note 538, at 95.
\textsuperscript{542} See Fisher, supra note 537, at 108; Rieke & Stutman, supra note 538, at 95.
\textsuperscript{543} See Bonnie & Slobogin, supra note 35, at 477-79. The example is embellished a bit for clarity, and, more importantly, the example is analyzed using the framework of storytelling theory, a framework not used by Professors Bonnie and Slobogin.
stories drawn from their own experiences. It seems implausible that Ms. B would believe Scott when one of his checks previously bounced, all the checks were made out to different payees, and he claimed that he lacked his own bank account. In the stories that jurors hear and experience, honest people do not behave like Scott.

Narrative fidelity is also weak because facts are clearly missing; surely there must have been conversations between Ms. B and her boyfriend about Scott’s request that should have been brought to light. Moreover, most jurors will assume that Scott is dishonest, and associating with such people violates fundamental social values favoring honest behavior. That Ms. B is the kind of person who would associate with Scott suggests that Ms. B is also dishonest, a defect of character, thus reflecting back on narrative coherence. Ms. B is, therefore, not someone to be believed.

An expert character witness—a psychiatrist—might testify, however, that Ms. B’s personality is characterized by an unusually high degree of passivity and dependency. These conclusions would be supported by the psychiatrist by recounting relevant portions of Ms. B’s life history, interviews, tests, and other data collected during her evaluation. Because Ms. B is dependent on others to satisfy her emotional needs, the psychiatrist would opine, she is compliant and generally avoids conflicts that threaten the stability of her emotional attachments. Consequently, in situations like this one, her dependence on her boyfriend and desire to please him would likely lead her to want to please her boyfriend’s good friend, Scott. As is characteristic of those with her personality traits, Ms. B would rely on denial and repression to keep out of her consciousness any anxiety-provoking thoughts that might threaten her emotional connection to her boyfriend. Thus, doubts about Scott’s honesty and, by implication, about her boyfriend’s character, would be repressed. Her unusual gullibility in situations of this kind would therefore explain her behavior.544

544. Professors Bonnie and Slobogin say in a footnote that they are “not sanguine” about the admissibility of expert character testimony. Id. at 481-82 n.162. They cannot literally mean what they say because much of their article is devoted to the contrary proposition: defending the admissibility of much testimony that meets this Article’s definition of “expert character testimony” (for example, they defend use of expert testimony regarding the personality of Ms. B. See id. at 477-79.). Consequently, although their discussion on this point is unclear, they seem instead to be suggesting two things. First, general testimony about a defendant’s “character” that is not related to the relevant events at the time of the crime—i.e., that is unrelated to the “situation”—is unhelpful. Second, expert opinions must address what the defendant’s mental state was,
The psychiatrist's testimony seeks to improve the narrative coherence and narrative fidelity of the defendant's testimony. Narrative coherence is improved by suggesting an alternative story line that may make sense when placed in light of the defendant's life experiences: she did what she did because her life story molded her into someone especially eager to please and to believe in those she loved. Although her reaction may seem extreme, jurors' own stories likely include actions motivated by the desire to help those close to them and the refusal to believe bad things about such people. Moreover, Ms. B can then be seen not as dishonest but rather as a sad dupe of her loved ones, and those associated with her loved ones, who have taken advantage of her.

Narrative fidelity is also improved by adding this type of information. The jury need not wonder whether there are missing facts, such as conversations between Ms. B and her boyfriend, for the jury is presented with other facts to fill the gap: facts regarding the experiences, feelings, and motivations of Ms. B that help to explain her behavior in a way that does not require the jury to condone dishonesty. While extraordinary gullibility may not be a trait to be admired, it is one that, unlike dishonesty, is more likely to evoke compassion and a greater willingness of the jury to judge her not by a "normal" standard of gullibility but rather as the abnormally gullible person that she is. Without the psychiatrist's evaluation, however, it is unlikely that the jury would give Ms. B's full story even passing consideration.

While storytelling theory, and the notion of "plausible alternatives to normal," thus help to illustrate the value of even "suppositional science," one more matter must be considered to craft a complete response to Professor Faigman's assault on the use of clinical testimony. That matter is this: Professor Faigman's position turns on the assumption that only "science" can be of value in the courtroom. But even accepting his characterization of clinical

not merely what kind of person he was, to ensure that we try "the case and not the man." See id. at 482 n.162 (quoting Thompson v. Church, 1 Root 312 (Conn. 1791)).

As to the first point, they are clearly correct. This Article has repeatedly emphasized the importance of the relationship between the person and her situation. As to the second point, however, they are mistaken. Expert character testimony is best understood as the use of knowledge of the interaction between personality and particular situations to estimate the probabilities that a defendant had a necessary mental state, not to opine what that mental state actually was. Claiming more of such testimony would mislead a jury about the testimony's value.

545. Faigman's criterion for scientific knowledge requires repeated attempts to determine the "falsifiability" of testable statements. Faigman, Social Science and Policy, supra note 407, at 1014-15.
testimony as "non-science"—as mere "art"—he is wrong, for there is great need for "art" in the courtroom, and juries are well equipped, when properly educated, to judge the quality of art as well as science.

c. Art versus Science.—

i. Why Art is Helpful.—The assault by a portion of the academic community on the forensic use of clinical judgment reflects what Professor Donald Schöen has criticized as the single-minded emphasis of professional education on the "science" rather than the "art" of professionalism. Professor Schöen bemoans this emphasis because he believes that artistry has much to offer that science cannot.

Science, Professor Schöen argues, reasons from standardized, general principles to the particular case. Selective inattention is therefore paid to that which does not fit within the general rules. Art, on the other hand, recognizes that much of what we know, we know implicitly from the way in which we solve individual problems; useful knowledge need not, therefore, be limited to applying general principles governing generalized situations. Indeed, this implicit knowledge can be studied and used by the artist to solve novel problems arising in individual cases. Such studies show us that professionals-as-artists thus approach each case—or clinicians each patient—as a "universe of one"; hypotheses are articulated by the artist and tested for usefulness. What works is what is chosen and followed, allowing for growth, change, consider-

546. See id. at 1085 (deriding Freudian psychologists as no better than would be "Dostoevskian psychologist[s]").

547. See Schöen, supra note 8, at vii ("[Educational] institutions [are] committed, for the most part, to a particular epistemology, a view of knowledge that fosters selective inattention to practical competence and professional artistry.").

548. Id. at 21-24. Professor Schöen describes the model that reveres science over art as the model of "Technical Rationality," a phrase that connotes a cold, mechanical, narrowly focused approach to problem solving. Such an approach seems to be the opposite of what legal decision making should embody. See supra text accompanying notes 487-513.

549. Schöen, supra note 8, at 44-45.

550. See id. at 140 ("Reflection-in-action in a unique case may be generalized to other cases, not by giving rise to general principles, but by contributing to the practitioner's repertoire of exemplary themes from which, in the subsequent cases of his practice, he may compose new variations.").

551. See id. at 49, 68 (discussing "tacit knowing" and noting that one who reflects on such knowledge is "not dependent on the categories of established theory and technique, but constructs a new theory of the unique case" (emphasis added)).

552. See id. at 105-28 (chapter entitled, "Psychotherapy: The Patient as a Universe of One.").
ation of otherwise overlooked options, and increased effectiveness in ways that a wholly scientific approach would not.553

The distaste of professional schools for the artistry that, in Professor Schön's view, should be at the heart of their endeavor,554 carries over to the rationalist critiques, like that of Professor Faigman, which seek to impose the standards of science in the courtroom. Despite its pretensions to the contrary,555 law is not a science but rather deals with the need for relatively quick, practical decision making in individual cases because judgments cannot be postponed while awaiting scientific confirmation.556 These decisions are made in a value-laden world of uncertainty. Thus, verdicts, being a practical project, do not and should not seek solely "truth," in some absolute, scientific sense of that term, but rather acceptability and stability, given the constraints and varying goals facing the decision-maker.557

Therefore, while science is not irrelevant—and, indeed, can do much to inform the legal process when truth is one of the competing goals—sometimes art can do in an acceptable fashion in the individual case what science cannot. This is especially true when there are powerful concerns other than truth, such as the political, moral, and cathartic goals of a criminal trial.558 We may, therefore, turn to art, in this instance, to clinical judgment with its uncertain scientific support, because the artistry serves cathartic goals or because, in our search for truth, artistry is the best means available. Indeed, the

553. See id. at 128-67 (comparing the experimental method with the professional artist's method, which latter method Schön describes as "reflection-in-action").
554. Id. at vii.
555. See id. at 28-29 (noting that despite law's "dubious basis in science," the dominant case-method of teaching fits the professional view of law as a "science").
556. See Charles Nesson, Agent Orange Meets the Blue Bus: Factfinding at the Frontier of Knowledge, 66 B.U. L. Rev. 521, 529-30 & n.26 (1986) ("[W]e ask juries to come to conclusions without insisting on or waiting for scientific demonstration.").
557. Id. at 521 ("An outcome is 'probable' if it best accomplishes a just and acceptable resolution of the dispute. Probability, as a legal concept in the law of proof, suggests wisdom, probity, and approbation—not favorable betting odds.").
558. SCHÖN, supra note 8, at 50 ("It is [the] . . . process of reflection-in-action which is central to the 'art' by which practitioners sometimes deal well with situations of uncertainty, instability, uniqueness, and value conflict."). Lest I be accused of inconsistency, I must acknowledge that in an earlier article I argued precisely that the courts had failed by ignoring the scientific bases of one particular type of evidence: dog-scenting evidence. See Taslitz, supra note 270, at 19. However, prosecution use of dog-scenting evidence to identify who committed a crime raises purely truth-based concerns—"Do we have the right culprit?"—and does not raise the moral and process-based concerns involved with expert character evidence. Moreover, I argue here that science does indeed have an important role to play in analyzing expert character evidence, but not the sole role in determining the question of admissibility.
professional psychological literature recognizes both the limitations and advantages of clinical artistry.559

Thus, the literature recommends a purely actuarial approach when the outcome to be predicted is objective and specific, and interest in the individual case is minimal.560 On the other hand, clinical judgment is recommended when (1) information is needed about areas or events for which no adequate tests are available, (2) rare, unusual events of a highly individualized nature are to be predicted or judged, or (3) the clinical judgments involve instances for which no statistical equations have been developed.561 These are precisely the situations likely to arise in criminal trials.

Even when actuarial formulae exist, their application may turn on the accurate collection and categorization of data by a clinician. Thus, certain characteristics of a subject may not be discoverable without extensive interviewing by a clinician, again requiring a form of clinical artistry.562

Furthermore, a case study is far more effective than a quantitative approach at "delineating the particulars of persons and their circumstances; conveying what individuals said, thought, felt, and did; representing the subjective meanings of actions and events; and developing content-specific or idiographic interpretations and explanations of individual behavior." While behavior can be directly observed and recorded, subjective meanings cannot, but rather must be inferred. The case study is the most effective means for so doing and for effectively communicating those meanings to others. Therefore, when mental state or subjective meaning, as opposed to proof of an act, is at issue in a criminal case, the case study is likely to be the most effective way of exploring and communicating that issue.

ii. "Good" Art versus "Bad".—To say that the clinical arts can and should play a role in legal decision making is not, however, to say that there are no standards for distinguishing good art from bad art. Indeed, some have argued that the distinction between art and

559. See infra notes 560-563 and accompanying text.
560. E.g., PHARES, supra note 420, at 269. An actuarial approach may also be advisable when there are especially strong reasons to be concerned about human judgmental error, bias, fatigue, or boredom. Id.
561. Id.
562. See id. at 266; NURCOMBE & GALLAGHER, supra note 430, at 610 ("The clinician is indispensable at the outset of the clinical process, when the world must be confronted; clues elicited, sifted and validated; [and] inferences drawn.").
563. RUNYAN, supra note 442, at 244.
science is itself misleading and that quantitative data and the rigid falsificationist standards of the experimental method should not set the boundaries for science.\textsuperscript{564} Thus Donald Campbell, long one of the foremost proponents of the experimental method, has noted that "the core of the scientific method is not experimentation per se, but rather the strategy connoted by the phrase ‘plausible rival hypotheses.'"\textsuperscript{565} This strategy, he continues, should include "a humanistic validity-seeking case study methodology that, while making no use of quantification or tests of significance, would still work on the same questions and share the same goals of knowledge [as do quantitative approaches]."\textsuperscript{566} Whether called art or science, therefore, there are ways of testing the value of intensive individual studies of a single case, which can and should be applied to gauging the quality of a clinician’s opinion and of a clinician’s reliance on studies that may not yet be fully experimentally validated.

What are these ways for distinguishing good and bad studies of individuals?\textsuperscript{567} A case or clinical study\textsuperscript{568} has the advantage of relying upon multiple sources of data, such as interviews, observations, psychological tests, and quantitative data, to draw conclusions.\textsuperscript{569}

An initial consideration, therefore, should be whether there has been an adequate reference to all potentially useful sources of data.

\textsuperscript{564} "Falsifiable standards" refers to the notion, most clearly articulated by Professor Karl Popper and embraced as to the social sciences by Professor David Faigman, that the most important hallmark of the scientific method is the ability to falsify propositions. Theories that have survived controlled efforts at falsification are thus successful scientific theories. See Faigman, \textit{Social Science and Policy}, supra note 407, at 1016-19. Social science theories not yet subjected to adequate attempts at controlled falsification are, therefore, merely "suppositional science." See \textit{id.} at 1052.


\textsuperscript{566} \textit{Id.} at 8.

\textsuperscript{567} Strictly speaking, measuring how "good" a clinical opinion is, that is, how likely it is to be accurate, is a reliability question. However, both the reliability and necessity aspects of the "art versus science" controversy are better understood if considered together. It is for that reason, therefore, that both aspects are addressed here under the single heading of "necessity."

\textsuperscript{568} The terminology is being used loosely here, for a "case" study refers to the presentation of detailed information about an individual life; an "idiographic approach" relies solely on data from the individual case; and "clinical judgment" and "statistical" or "actuarial" judgment are ways of analyzing data, regardless of its source. See Runyan, \textit{supra} note 442, at 186. These distinctions are not key in this Article, for what matters is that expert character testimony may, and usually does, include data both from the individual case and, either explicitly or implicitly, from groups, and may analyze that data by use of either clinical judgment or statistical methods or both. See \textit{supra} text accompanying notes 121-136, 406-457.

\textsuperscript{569} \textit{Yin, supra} note 565, at 96.
and whether those sources have been thoroughly explored, that is, whether there is an adequate "data base." \cite{570} Database records must carefully maintain a "chain of evidence" by identifying the sources of the data, the evidence revealed by each source, the circumstances under which the evidence was obtained, and the consistency of these circumstances with the question being studied. \cite{571} "Converging lines of inquiry" are then sought in an effort to find corroboration of working hypotheses among the various data sources. \cite{572}

There may be a variety of potential hypotheses that explain the evidence generated. The pivotal step is to determine how a choice can be made among these hypotheses. One author has suggested the following general criteria for evaluating alternative hypotheses:

1. their logical soundness,
2. their comprehensiveness in accounting for a number of puzzling aspects of the event in question,
3. their survival of tests of attempted falsification, such as tests of derived predictions or retrodictions,
4. their consistency with the full range of available relevant evidence,
5. their support from above, or their consistency with more general knowledge about human functioning or about the person in question,
6. their credibility relative to other explanatory hypotheses. \cite{573}

These or similar factors can be approached in a quasi-judicial fashion, in which issues are clearly articulated, evidence is subjected to critical examination by others, the sufficiency of available empirical data both to support and to refute relevant claims is considered, and a logical connection is demanded between a claim and the evidence offered in support of it. \cite{574}

The point is that standards do exist for judging "art" as well as science, standards that can be articulated to, and understood by, a jury. The strong need for art, therefore, is not to be feared but rather to be embraced by the criminal justice system.

d. Availability of Less Objectionable Alternatives.—The need for expert character evidence cannot be fully understood without considering whether there are alternative types of evidence of similar or

\begin{footnotes}
\footnote{570. See id. at 98-101 (noting that an adequate data base increases the reliability of the case study).}
\footnote{571. Id. at 102-03.}
\footnote{572. Id. at 97.}
\footnote{573. Runyan, supra note 442, at 47.}
\footnote{574. See id. at 160-63.}
\end{footnotes}
greater probative value that may present less danger of prejudice while achieving the same results.\textsuperscript{575} The most obvious alternatives are evidence directed to insanity, diminished capacity, and more traditional forms of character evidence.

Insanity\textsuperscript{576} and diminished capacity\textsuperscript{577} can profitably be considered together because they share some common characteristics. Both require the jury to engage in the difficult, uncertain, and confusing inquiry of whether the defendant suffered from a "mental disease or defect."\textsuperscript{578} Character evidence may involve such testimony, but may also involve testimony regarding relatively "healthy" persons.\textsuperscript{579} In any event, when expert character evidence is admitted, the health or illness of the accused is not necessarily of concern to a jury.\textsuperscript{580}

Insanity and diminished capacity, unlike expert character testimony, require psychologists to opine on an accused's capacity to form the necessary mental state.\textsuperscript{581} This is especially problematic

\textsuperscript{575} See McCord, Profiles, supra note 119, at 96, 98 (illustrating this point through an example of potentially different evidence that could be presented during a child sexual abuse case).

\textsuperscript{576} The classic test for insanity is the M'Naghten rule, which states that a defendant cannot be convicted if, at the time the act was committed, the defendant was laboring under such a defect of reason from a disease of the mind as not to know the nature and quality of the act or that the act was wrong. M'Naghten Case, 8 Eng. Rep. 718 (1843). See also Dressler, supra note 4, § 25.04[C][1][a] at 299. Some jurisdictions will also find a defendant insane if she acted under an "irresistible impulse." \textit{Id.} at 301. See, e.g., State v. Hartley, 565 P.2d 658, 660 (N.M. 1978) (expanding the M'Naghten rule to provide that "if by reason of disease of the mind, defendant has been deprived of . . . his will which would enable him to prevent himself from doing the act, then he cannot be found guilty"); Thompson v. Commonwealth, 70 S.E.2d 284, 292 (Va. 1952) (recognizing that the irresistible impulse doctrine is applicable when a defendant's "mind has become so impaired by disease that he is totally deprived of the mental power to control or restrain his act").

\textsuperscript{577} "Diminished capacity," unlike insanity, requires proof that a defendant lacked the mental capacity necessary to commit the crime. Some jurisdictions extend the defense to all crimes, but most limit the defense to either some or all specific intent crimes. See Dressler, supra note 4, § 26.02[B][1].

\textsuperscript{578} See Peter W. Low et al., The Trial of John W. Hinckley, Jr.: A Case Study in the Insanity Defense 20 (1986) (explaining that the insanity defense requires proof of a mental disease or defect); Bonnie & Slobogin, supra note 35, at 474 (noting a similar requirement for the diminished capacity defense).

\textsuperscript{579} See Bonnie & Slobogin, supra note 35, at 479; supra text accompanying notes 543-545 (illustrating this through the case of Ms. B).

\textsuperscript{580} See Low et al., supra note 578 (generally tracing the confusion over disease diagnosis in the trial of John Hinckley, Jr.).

\textsuperscript{581} Both the insanity and diminished capacity defenses may require an inquiry into "capacity." Dressler, supra note 4, § 25.04[3][a], § 26.02[3] & n.10. Some versions of the insanity defense, however, require merely that the defendant lacked "substantial capacity," instead of lacking any capacity. Low et al., supra note 578, at 18-20. Even in
for diminished capacity because "[o]nly in connection with the concepts of 'deliberation' and 'premeditation' is it clinically conceivable that a legally sane defendant would nonetheless lack the mental capacity to perform the requisite cognitive functions."582 Additionally, insanity is rarely used and more rarely successful, partly because it requires proof of things unlikely to happen or extraordinarily difficult to prove.583

Character testimony, on the other hand, involves the expert in a more realistic and more broadly useful universe of probabilities, through an inquiry into tendencies and predispositions instead of absolutes. Such testimony probes into the area of probable thoughts and behavior, not the nebulous doctrine of insanity, in which law and morality overtly blend, where truly bizarre thinking is standard fare, and where a defendant's actual state of mind, not her likely one, is the subject of the testimony.584

More traditional forms of character evidence—specifically, reputation, lay opinion, and prior good and bad acts—while not suffering from the same flaws as insanity and diminished capacity—suffer from other defects that make them less effective than expert character testimony. Reputation, for example, more properly concerns the "shadow" a person casts on the community rather than the actual character of the person, two sharply different notions.585 Testimony as to reputation is typically "bloodless, ritualistic and a bit dull."586 There is, therefore, little assurance that reputation is based on trustworthy data reflecting an individual's true character and little opportunity to test the bases for reputation effectively.587 For example, an inquiry into what particular informants told a reputa-
tion witness is often excluded under the hearsay rule.\textsuperscript{588} This is so even though "reputation" is merely the sum of individual informants' opinions. Furthermore, the sources of the gossip are unavailable for cross-examination to test the bases for their views of the defendant. Finally, even if particular out-of-court statements were admissible, community gossip is not likely to connect a person's traits to the particular situations that evoke the traits at issue in the way that expert character evidence can.

On the other hand, lay opinion—the "warm affectionate testimony" of friends, neighbors, co-workers, and family—might be of greater impact and perhaps of greater probative value than reputation.\textsuperscript{589} But lay opinion suffers from the defects of potential bias and the difficulty of a layperson's articulating in any precise, specific fashion data to support his favorable opinion of the accused.\textsuperscript{590} A jury will be left with anecdotes and conclusions, rather than useful information, and will likely rest its judgment primarily on the witness's demeanor.

This criticism should not be overstated, of course, for some have concluded that laypersons are often effective judges of character. This is particularly true when a defendant is well known to the laypersons and they have had an opportunity to observe the defendant across a wide range of situations.\textsuperscript{591} Nevertheless, generalized

\begin{itemize}
  \item sources of information because impeaching the witness by an inquiry into what particular informants stated would violate the hearsay rule.

  \textsuperscript{588} \textit{Id.} Although Federal Rule of Evidence 803(21) creates an exception to the hearsay rule for "reputation of a person's character among associates or in the community," that exception is often construed as overcoming the hearsay objection for reputation as a summary of the out-of-court statements of many others in the community, not as an exception for individual out-of-court statements. \textit{See} John C. O'Brien \& Roger L. Goldman, \textit{Federal Criminal Trial Evidence} 198, 422 (1989). \textit{See also} United States v. Webster, 649 F.2d 346 (1981) (stating that reputation evidence may be given only by a person familiar with the defendant's reputation and competent to speak for the community).

  \textsuperscript{589} Put any one of us on trial for a false charge, and ask him whether he would not rather invoke in his vindication, as Lord Kenyon said, "the warm, affectionate testimony" of those few whose long intimacy and trust have made them ready to demonstrate their faith to the jury, than any amount of colorful assertions about reputation.

  \textsuperscript{590} \textit{See infra} text accompanying notes 617-640.

  \textsuperscript{591} \textit{See} Davies, \textit{supra} note 360, at 522 ("One study shows that it takes 1,440 direct measurements of behavior, made over a period of eight weeks, to predict the average behavior of a subject as reliably as the personality judgments of three knowledgeable informants.").
\end{itemize}
trait descriptions\textsuperscript{592} are still of much less value than descriptions that refer to specific situations in which a person has demonstrated the trait in the past.\textsuperscript{593} Indeed, this latter point is particularly important because much of the effectiveness of predicting behavior in everyday life is due more to the situational component than to knowledge of the "character" component of people's lives; thus, professors can be expected to be professorial, dictators dictatorial and servants servile, because that is what their roles demand.\textsuperscript{594} An expert will be more aware of the respective contributions of traits and situations and hence of the value of the data upon which to base an opinion in the case before the jury.

A third form of traditional character evidence—specific instances of prior conduct—may also be used to prove state of mind, but not to prove the criminal act.\textsuperscript{595} Yet, a specific act may not be representative of a defendant's normal character or way of functioning.\textsuperscript{596} Jurors are nevertheless likely to treat a single act as demonstrating what an actor ordinarily does.\textsuperscript{597} Of course, when there are a larger number of specific acts—a wider sample of behavior brought to the jury's attention—and when those acts are quite similar to the one with which the defendant is charged, the probative value of prior acts is increased.\textsuperscript{598} Even then, however, the impact

\textsuperscript{592} Generalized trait descriptors are the kind most in accord with the common speech of laypersons. For example, most would say "John is a peaceful person," rather than "John is peaceful in classes and the school lunchroom."

\textsuperscript{593} Davies, supra note 360, at 521-23.

\textsuperscript{594} Ross & Nisbett, supra note 365, at 150.

\textsuperscript{595} See Fed. R. Evid. 404(b) (barring admissibility of other crimes, wrongs, or acts to prove character to show action in conformity therewith, but admitting such acts for "other purposes," such as proof of intent). Rule 405 also permits proof of specific instances of conduct in cases when "character or a trait of character of a person is an essential element of a charge, claim, or defense," Fed. R. Evid. 405(b), but such a use of character is entered as direct, not circumstantial, evidence. Fed. R. Evid. 405 advisory committee's note. Inquiry into specific instances of conduct is also permitted for impeachment purposes when cross-examining a reputation or opinion witness as to a defendant's character. Fed. R. Evid. 405(a). See also Fed. R. Evid. 608(b) (permitting, under certain circumstances, testimony concerning specific acts, but only if offered to prove the character of a witness for truthfulness or untruthfulness).

\textsuperscript{596} See Mendez, supra note 153, at 1046-50 (explaining the tendency of people, when predicting behavior, to generalize from one act).

\textsuperscript{597} Id. at 1048.

\textsuperscript{598} See Davies, supra note 360, at 519-20, 535-36 (recommending removal of the current per se restriction on the prosecution's first use of prior bad acts to show conduct and its replacement with a balancing test based on the specificity of the subject's prior conduct, the similarity between that conduct and its surrounding circumstances to the conduct in question, the number of provable instances of prior similar conduct, the period over which the prior conduct occurred, and whether the conduct was inherently likely to be repeated).
on the jury may be inappropriate because of the "halo effect," the tendency of people to judge others on the basis of one outstanding "good" or "bad" quality. This effect may tempt jurors to judge a defendant based on their generalized perceptions of her rather than on the evidence as to her actions or thoughts on the occasion in question. This may be true even when the perception is based on a single prior act or a small number of acts.

Psychological character testimony may also result in the revelation of certain prior acts. Those acts, however, will not be viewed in isolation. Unlike specific instances of prior conduct, they will be viewed as part of a broader database designed to give a jury a fair picture of how the actor usually thinks or behaves in similar situations. Because effective experts are trained to behave like detectives, they will seek additional data to expand their knowledge of the nature and frequency of prior acts so that they can make a more informed judgment. This incentive may be missing for the layperson. Moreover, with expert character witnesses, the precise bases, limits, and value of an opinion can be extensively probed on both direct and cross-examination. Such extensive information will better enable a jury to decide how much it should rely on the opinion, a decision impossible to make were the jury shown only bare reports of bad acts alone.

One specific type of prior acts evidence—prior criminal convictions—should be separately addressed. Under the Federal Rules of Evidence, prior criminal convictions can be used to impeach any witness if the crime involved dishonesty or false statements. If the crime did not involve dishonesty or false statement, a witness may be impeached only if the crime was punishable by death or imprisonment in excess of one year, and if testimony about the conviction

599. Mendez, supra note 153, at 1047.
600. See id. (discussing jurors' propensity to judge people based on a "unity of personality").
601. See supra text accompanying notes 142-243.
602. Prior acts may be explored on cross-examination of a lay character witness to impeach that witness's opinions, but not to prove the conduct at issue in the case before the court. See Weissenberger, supra note 18, at 98-99. Davies argues that when prior acts survive a balancing test focused on the number and similarity of the prior acts to the conduct in question, they should be admitted as substantive proof of conduct in the present case. See Davies, supra note 360, at 535. Because an expert is likely to have uncovered even more acts identified as situationally specific to the conduct in question than would a lay witness, Davies's approach has even more merit as a justification of the admissibility of prior acts revealed in cross-examining expert character witnesses.
is determined by the court to be more probative than prejudicial. 604

The latter part of this rule turns on a character judgment unsupported by modern psychological theory. No theory suggests that a witness will lie to the court merely because that witness previously was convicted of a crime unrelated to lying. 605 Yet the rules assume that someone convicted of an offense punishable by death or imprisonment for more than one year is indeed the kind of person who is more likely to lie. This assumption ignores the need for more-narrowly defined traits and the powerful role of the situation in determining human behavior. 606 Aggravating the fallacy of these assumptions is the possibility that the jury may be asked to draw broad conclusions about honesty based on a narrow sample of one or two prior convictions. 607 This ignores the principles of aggregation discussed above. 608 Indeed, even the use of convictions for dishonesty or false statements, without further limitations on admissibility, may be unwise. Honesty is a trait very much influenced by situational factors; a willingness to lie in one situation, therefore, says little or nothing about the propensity to lie in a different situation. 609

Finally, a comparison must be made between habit evidence and psychological character evidence. Habit evidence is admissible to prove that a person’s conduct was in conformity with the habit displayed on a particular occasion. 610 Evidence of habit, therefore, is not subject to Federal Rule of Evidence 404’s character evidence ban. 611 While character concerns one’s general disposition or one’s

---

604. Id. When the accused’s conviction is offered, the court must determine that the probative value of the evidence outweighs the prejudice to the accused. Id. When the conviction concerns someone other than the accused, the test is the standard pragmatic relevancy balancing test: whether probative value is substantially outweighed by countervailing considerations. See id.; Fed. R. Evid. 403.

605. See Davies, supra note 360, at 520-21.

606. See supra notes 142-163 and accompanying text; Davies, supra note 360, at 520-21 (describing the underlying assumption that a prior conviction demonstrates a “readiness to do evil,” from which a jury is meant to infer a readiness to lie in the specific case (quoting Gertz v. Fitchburg R.R., 137 Mass. 77, 78 (1884))).

607. See Davies, supra note 360, at 520-21.

608. See supra text accompanying notes 363-374.

609. Davies, supra note 360, at 521. Davies’s suggested cure is to admit evidence of the prior bad acts if placed in the context of the situation that produced the acts. To avoid jury prejudice from learning of an official condemnation of those acts, however, Davies would delete reference to the fact of conviction unless necessary to rebut a defendant’s denial that the acts took place. Id. at 536-37.


disposition in respect to a general trait, habit is a more specific concept describing one’s general response to a repeated specific situation. But if, as many psychologists contend, personality traits should often be more carefully defined in terms of likely responses in certain situations, how can “traits” and “habits” be distinguished?

From an evidentiary perspective, this distinction may be made by recognizing the primary justification for the court’s greater willingness to accept habit evidence than character evidence: the assumption that habit involves more than a routine practice—that a habit must also be an unconscious, mechanical, and at least “semi-automatic” response to a very particular kind of situation. So viewed, character differs from habit in that character concerns general dispositions to behave in a certain way, a quality thought to have less predictive power in a given situation than the automatic reaction that is habit. Moreover, from the psychologist’s perspective, trait often remains a more general concept than habit; for example, the trait of cleanliness might in part be viewed as synthesizing the habits of brushing one’s hair and teeth, washing one’s hands, cleaning one’s nails, and wearing clean clothing.

While psychological character testimony may, therefore, be based partly on observations of habits, such testimony should be governed by character evidence rules alone. The greater resulting difficulty in admitting psychological evidence of this nature will not diminish the potential usefulness of character evidence over habit. For example, it is highly unlikely that a habit for killing someone can be demonstrated, but a “character” for violence or for peacefulness, at least in certain ranges of situations, can. Moreover, habit’s usefulness is probably limited to proving conduct, while, as defined in this Article, character may also be used to prove state of mind, a central question in much criminal litigation.

612. FED. R. EVID. 406 advisory committee’s note.
613. See id.
614. See id.
615. See HERGENHAHN, supra note 144, at 184; see also FED. R. EVID. 406 advisory committee’s note (“Character may be thought of as the sum of one’s habits though doubtless it is more than this.” (quoting MCCORMICK ET AL., supra note 16, § 195, at n.7)).
616. “Habit” is by definition an observable “response.” See FED. R. EVID. 406 advisory committee’s note. Even were this not so, it is hard to prove in any meaningful way a “habit” of thought that is so situationally specific, frequent, and automatic as otherwise to fit Rule 406’s definition of “habit.”
3. Understandability.—Understandability gauges the jurors’ ability both to comprehend testimony and to give it proper weight.617 The law of evidence’s long-standing hostility toward much character evidence is based on the assumption that such evidence will not be properly understood, will be overvalued, or will lead juries to convict because of a defendant’s character rather than what that defendant thought or did.618

Psychological research suggests that there is good reason to worry about jurors’ overvaluing character evidence. Psychologists express this concern by describing the well-documented “fundamental attribution error”: the tendency of people to inflate the importance of personality traits while failing to recognize the importance of situational factors in prompting behavior.619 People often overlook important situational factors and “infer dispositions from behavior that is manifestly situationally produced.”620

A classic illustration is Darley and Batson’s “Good Samaritan” experiment, in which some students at a religious seminary were told that they were late for a talk while others were told that they had plenty of time.621 Both groups passed by a man slumped in a doorway, groaning and coughing.622 Only ten percent of the “hurrying” seminarians stopped to help, while sixty-three percent of those not hurrying stopped to help.623 This suggests that the situational factor of lateness was the most important factor affecting whether the seminarians would be “good samaritans.”624 Yet, in a similar experiment, when laypersons were asked how they thought people would behave, the subjects predicted that the great majority of seminary students would stop to help and that being in a hurry would make no difference at all.625 They believed very simply that “altruistic people” would help while “selfish people” would not, regardless of how much time they had on their hands.626

Related concepts raise additional reasons for concern. Notably, laypeople overestimate the ability to predict from one observation

617. McCord, Profiles, supra note 119, at 100.
618. Davies, supra note 360, at 523-25.
620. Id. at 126.
621. See id. at 48-49.
622. Id. at 49.
623. Id.
624. Id.
625. Id. at 131 (describing a follow-up experiment conducted by Pietromonaco and Nisbett).
626. Id.
what will happen on another occasion, showing little understanding that predispositions address average behavior only, while individual instances of behavior remain highly variable.\textsuperscript{627} Moreover, laypeople show little awareness of the advantages of aggregating many behavioral measures to craft better predictors of average behavior.\textsuperscript{628} This suggests in turn that people are generally willing to make confident, disposition-based predictions, based upon limited data.\textsuperscript{629}

Furthermore, because they ignore situational factors, laypeople are much more willing to overgeneralize: to assume wide predictive power of a trait across a broad range of situations when the true range may indeed be quite narrow.\textsuperscript{630} Additionally, laypeople often ignore the importance of base rates as central to accurate predictions of behavior.\textsuperscript{631}

Each of these tendencies, however, should be of greater concern with lay character testimony than with expert character testimony, because expert testimony can counteract these tendencies in a number of ways. First, competent expert character testimony will, by definition, emphasize the interaction of traits and situations. Preliminary research conducted in this area suggests that expert character testimony helps to focus juries on the importance of situational factors in eliciting behavior and in offering opportunities for traits to manifest themselves.\textsuperscript{632}

Second, expert character testimony educates jurors about a variety of matters that will help them place the testimony in proper perspective and make informed judgments. Most importantly, that education should emphasize the need for a wide sample of behavior—aggregation—to gauge traits accurately, and the usefulness of traits as predictors of average behavior only. Preliminary research on the effect of educating juries on "social frameworks"—comparisons between a participant in litigation and a larger social grouping, similar to what has been called "group character evidence" in this

\textsuperscript{627} See id. at 122-23.
\textsuperscript{628} See id. at 123-24.
\textsuperscript{629} Id. at 124.
\textsuperscript{630} See id. at 129-30. On the other hand, at least one study suggests that when people are given information about specific acts in situational settings, rather than in the form of a global trait descriptor, they may not always experience fundamental attribution error. Davies, supra note 360, at 531-32 & n.159.
\textsuperscript{631} Ross & Nisbett, supra note 365, at 134-36.
\textsuperscript{632} Vidmar & Schuller, supra note 135, at 154, 160 (reporting experimental data that suggests that both battered woman and rape trauma syndrome testimony draw the jury's attention to contextual and situational factors).
Article suggests that, at least under certain conditions, expert education of juries does affect their decision making without “overawing” them.

Third, expert testimony can alert jurors to their own beliefs regarding personality traits and point out logical errors made in acting on those beliefs. Again, while the evidence is sparse and tentative, there is some support for the notion that jurors can improve the accuracy of their decisions if alerted to their own biases and decision-making flaws. A caution is in order: improving the quality of juror decision making does not necessarily mean eliminating the effect of juror reasoning errors and preconceptions; rather, it may merely mean reducing that effect.

Fourth, the tendency of jurors to make “snap judgments” about character based on limited information can only be counterbalanced by giving them more information. For a psychological character expert’s testimony not to be undervalued, the expert must relate broad descriptions of group behavior to the behavior of the specific defendant. So doing requires the injection of clinical judgment based on the wide database discussed earlier—large samples of the defendant’s prior behavior; descriptions by family members, friends, and co-workers; results of psychological tests, interviews, observations, and so on. Jurors will then have a much larger and


634. See, e.g., Vidmar & Schuller, supra note 135, at 154 ("The introduction of expert testimony regarding the battered woman syndrome appears to have a salutary effect on the more prejudicial beliefs, but the primary effect seems to be to shift jurors toward the lesser included offense of manslaughter rather than toward a verdict of not guilty . . . ."); ABA, JURY COMPREHENSION, supra note 535, app. 10, at 9 (noting that a study on the impact of probability data in the form of expert scientific testimony in a rape trial concluded that "jurors may make considerable use of the expert testimony, even though they may not spend a great deal of time discussing it and when asked they may report it had little impact on their decision . . . .").

635. See, e.g., NISBETT & ROSS, supra note 344, at 191 (discussing experiments in which subjects confronted with information discrediting certain faulty beliefs nevertheless persevered in those beliefs; that tendency to persevere, however, was almost totally eliminated when those subjects were given an explicit theoretical explanation of the mental processes that cause them to hang on to their flawed beliefs).

636. See ROSS & NISBETT, supra note 365, at 131-32 (noting that telling subjects about the results of the “Good Samaritan” experiment, see supra text accompanying notes 621-626, did have an impact on their estimate of the effects of hurrying; but the effect was a mere 18% difference, much less than the 53% difference in fact observed by Darley and Batson).

637. See sources cited supra note 328.
fairer sampling—a better database—upon which to base their character judgments.

Finally, it must be emphasized that in many cases, the potential for jury prejudice may be relatively small when weighed against the potential gains offered by expert psychological character testimony. This is particularly true when the subject matter of the expert testimony is very close to the jurors’ common experience, because they are then more likely to understand the testimony fully. But the magnitude of the prejudicial effects will vary, sometimes dramatically, based on the specific type of expert character evidence offered and the facts of the specific case. The availability of literature addressing the value of particular types of character evidence—a literature that is now growing—will improve the quality of lawyers’ preparation and cross-examination and, therefore, improve jurors’ ability to accord the testimony its proper weight.

4. Weighing the Balance.—Having identified the primary areas for analysis, the question remains: how should a judge weigh these various factors? The answer will, of course, vary with the precise type of evidence involved and the facts of specific cases, but several general observations can be made to guide the weighing process.

The most important observation is that the moral and social psychological justifications for admitting psychological character evidence, and the related cathartic functions served by such evidence, must be an important part of the weighing process though we should not blindly ignore truth-seeking as a predominant value. Attention to goals other than truth-seeking, however, suggests several consequences.

Notably, a court should not view its decision as “either/or”: to admit or exclude. Instead, the court should take an activist role in pursuit of a third alternative: full disclosure of all relevant information to both the court and the jury regarding the strengths and weaknesses of the evidence, presented in a fashion likely to reduce jury misuse and increase accurate appreciation of the value of the evidence offered. Recent scholarship has recognized that in-

638. See McCord, Profiles, supra note 119, at 100 (suggesting that the more jurors are able to relate expert testimony to their common understanding, the more likely they will afford the testimony its proper weight).
639. See id. at 100-01.
640. See id. at 100.
641. The activist role suggested here is admittedly time-consuming. However, courts should be most receptive to expert character testimony when a defendant claims to have unusual or unique traits. Expert character evidence should usually be rejected in run-of-
creasing the quantity and quality of evidence is often an implicit goal of much evidence law. There is a strong justification for a trial court to make that goal explicit and to take an activist role—not relying solely on the parties—to ensure that the best quality and quantity of data regarding the value of the psychological character evidence are presented to a jury. Only if, after making such an effort, a court remains convinced that such evidence cannot be helpful to the jury, should it be excluded.

An activist judge would, therefore, require "Brandeis briefs" concerning the value and likely impact of the evidence before the court; would carefully examine the extent and status of discovery regarding the evidence and issue any orders necessary to ensure full and fair disclosure; would insist on detailed proposed jury instructions regarding the evidence; and would require that briefs and oral argument carefully address each of the concerns identified in this Article. Admittedly, in many jurisdictions discovery rules in criminal cases are so limited that this task will be difficult. However, changes are under way. For example, a proposed amendment to Federal Rule of Criminal Procedure 16 would require the prosecution to disclose to a defendant any expert evidence that the prosecution will offer at trial. Such disclosure would be required to be in the form of a written report prepared and signed by the witness, the-mill cases, thus limiting the number of cases in which an activist judge must invest her time. Moreover, the value of a certain technique need not be re-examined in every case if the technique was thoroughly analyzed before, although particular applications of that technique should still be open to dispute. Furthermore, the pretrial screening methods discussed throughout this Article, while time-consuming, should at least prevent jury trial time from being wasted on frivolous claims. Finally, some significant investment of time is simply unavoidable under our criminal law. The system must choose: either abandon morality-based judgments or sometimes pay the price of impaired efficiency to enhance higher goals. See Tyler, supra note 341, at 110 (arguing that citizens will accept delays from increased caseloads in the courts if they believe that there is fairness in trial procedures and in the allocation of benefits and burdens of litigation).


643. A "Brandeis brief" is an extensive brief containing economic and social survey material along with legal principles. The name is taken from the late Supreme Court Associate Justice Louis D. Brandeis. See, e.g., Muller v. Oregon, 208 U.S. 412, 419-20 (1908) ("In the brief filed by Mr. Louis D. Brandeis, [then a practicing attorney] . . . is a very copious collection of all these matters . . . .").

including a complete statement of all opinions to be expressed, the basis and reasons therefor, the data or other information relied upon in forming such opinions, any exhibits to be used as a summary of or support for such opinions, and the qualifications of the witness.\textsuperscript{645} The proposal also would require disclosure of "any information which might be recognized as a legitimate basis for an opinion under Federal Rule of Evidence 703, including opinions of other experts."\textsuperscript{646} The prosecution would be entitled to similar reciprocal disclosure by the defense.\textsuperscript{647} Recent scholarship has argued as well that court-appointed experts can be more extensively used under Rule 706, which permits the scheduling of expert depositions, and has argued that, in any event, courts have inherent power to order depositions of any expert witnesses.\textsuperscript{648} The tools for active judicial control of effective discovery are thus becoming available. Furthermore, judicial activism and more effective discovery rules are likely to develop in tandem, as one lends impetus to the other.

An activist trial judge would also closely scrutinize an expert's qualifications, not simply to determine whether the expert is qualified to make the relevant judgment, but also to determine whether the expert is adequately qualified to inform the jury fully and fairly. Such an expert must have both practical experience and substantial theoretical knowledge, and any expert who fails to meet the necessary background and abilities to educate the jury should be barred from testifying.

Additionally, an activist judge must weed out highly generalized trait descriptions that are likely to mislead juries. Descriptions or explanations that fully outline the relationship between the posited traits and the particular type of situation that brought a defendant into court should be encouraged instead. Testimony must also be qualified in terms of cautious probabilities, not certainties, and of the likely impact of different posited versions of the relevant situation on a defendant's thoughts and behavior. This last point is especially important because the central dispute in a criminal case often concerns the situation indeed faced by the defendant.

The importance of the cathartic function has implications for the standard of proof to be applied in determining whether the probative value of the proffered evidence is substantially outweighed by

\textsuperscript{645} Id.

\textsuperscript{646} Id. at 65 (emphasis added).

\textsuperscript{647} Id. at 63.

countervailing concerns. Professor McCord has argued, based on constitutional grounds, that there is a strong justification for imposing a higher burden of proof on the prosecution than on the defense regarding the reliability of nontraditional psychological evidence.  

This Article reaches a similar conclusion but on different, non-constitutional, grounds; specifically that, as discussed earlier, the cathartic values favoring admissibility are substantially greater for psychological character evidence offered by the defense than by the prosecution. This Article recasts Professor McCord's focus from the preliminary question of reliability to the preliminary question of whether probative value is substantially outweighed by other concerns. While reliability is one preliminary fact to be determined in the admissibility decision, many other facts—for example, jury impact, availability of alternatives, qualifications of the expert—must also be determined to find the "ultimate preliminary fact"; the degree to which probative value is or is not outweighed by other concerns. The critical question, therefore, must ultimately be this: How convinced must a judge be that the test of pragmatic relevancy has been passed?

Recognizing, as did Professor McCord, that the answer to this question may differ for defense and prosecution, the question must be posed separately for each. Professor McCord suggests that a preponderance-of-the-evidence standard is adequate as to the prosecution, rejecting Professor Giannelli's argument that a beyond-a-reasonable-doubt test is required. Professor McCord would, under his own version of the pragmatic relevancy balancing test, require proof of a higher degree of reliability as the issues in question grow "more important"—that is, the more likely the issue is to be dispositive of the case. For such issues, he argues that only "highly reliable" evidence should be admitted. Because such evidence must be "highly reliable," he argues that there is no need for the high standard of proof beyond a reasonable doubt.

But, as applied to expert character evidence, there are several flaws in Professor McCord's recommendation. First, he offers as an

649. See McCord, Profiles, supra note 119, at 105-07.
650. Many of the constitutional issues raised by psychological character evidence will be addressed in an upcoming article.
651. See supra text accompanying notes 391-345.
652. See supra text accompanying notes 347-354.
653. See McCord, Profiles, supra note 119, at 106.
654. Id. at 101.
655. See id. at 101-02, 106.
656. Id. at 106.
example of an issue that would be "decisive" an expert's testimony that an alleged child sexual abuse victim was indeed abused.\textsuperscript{657} However, as defined here, expert character evidence would never include an opinion that an act or a thought occurred. Rather, the opinion would be limited to identification of a defendant's trait and, in the usual case, use of that trait as circumstantial evidence of the likelihood that an element of a crime or defense existed.\textsuperscript{658} Expert character evidence would thus rarely, if ever, be "decisive" in the way envisioned by Professor McCord.

Second, the importance-reliability relationship suggested by Professor McCord offers judges even less guidance in exercising their discretion than is now true, for conceivably there can be an infinite range of levels of required reliability for different levels of importance. A single, familiar standard of proof is more likely to give courts clearer guidance.

Third, because of the need to give parties a strong incentive to be fully prepared and provide the court with as much information as possible—a need clearly recognized by Professor McCord but apparently not viewed by him as relevant to choosing the appropriate level of proof\textsuperscript{659}—a high standard should be adopted to motivate both prosecutors and defense counsel to prepare thoroughly.\textsuperscript{660}

Anything less than a preponderance-of-the-evidence standard is unlikely to do this. This Article, therefore, suggests that a preponderance standard should be the minimum applied to either side, and rejects Professor McCord's view that a "reasonable possibility" standard would be adequate as to the defense.\textsuperscript{661} However, if we recognize the need to impose a higher standard on the prosecution for first use of expert character evidence—because of the lesser prosecution claim to promoting cathartic values—the prosecution standard should be greater than a preponderance of the evidence. But this does not necessarily mean a standard of proof beyond a reasonable doubt.

Indeed, one of the reasons articulated by Professor Giannelli in first suggesting the beyond-a-reasonable-doubt standard for the prosecution is that scientific evidence often may result in erroneous

\textsuperscript{657} \textit{Id.} at 101.
\textsuperscript{658} \textit{See supra} text accompanying notes 16-24.
\textsuperscript{659} \textit{See McCord, Profiles, supra} note 119, at 95-98, 106-07 (suggesting that the level of proof should be decided by an estimate of how likely it is that the proposed testimony will assist the jury).
\textsuperscript{660} \textit{See Taslitz, supra} note 270, at 123 n.646 (discussing a similar point).
\textsuperscript{661} McCord, \textit{Profiles, supra} note 119, at 105-07.
verdicts, either because the issue involved is dispositive of the case or because the evidence is likely to overawe the jury. As just noted, rarely will expert character evidence be dispositive of a case, and, while jury impact must be assessed in every case, it is unlikely that psychological character evidence will "overawe" the jury in most cases, and much of any undue impact can be significantly moderated by an activist trial judge. The highest standard of proof, therefore, is probably not needed. The proposal here, therefore, is that the standard of proof for prosecution first-use of psychological character evidence should be "clear and convincing evidence," a standard high enough to recognize the reduced cathartic justification for prosecution-proffered character evidence and to offer a strong incentive for thorough prosecution preparation, but not so high as unduly to burden reasonable prosecution efforts to persuade the jury. Remember, however, that when a prosecutor seeks to make first use of expert character evidence to prove an act, Federal Rule of Evidence 404 will act as a flat bar. The "clear and


663. See supra text accompanying notes 293-308, 642-649.

664. See supra text accompanying notes 244-265. A similar flat bar arguably should apply to prosecution first-use of one type of expert character evidence—"group" character evidence—even when offered to prove mental state. The primary argument would be that when offered by the prosecution, group character evidence is so inflammatory that a jury is likely to ignore all evidence in the defendant's favor; ignore the need ultimately to judge the defendant's actual, individual mental state; and instead convict him merely for his apparent membership in a group.

For example, the prosecution might, to prove a defendant's awareness of lack of consent in a rape prosecution, offer evidence that the defendant was a "typical, violent rapist." Therefore, he thought as violent rapists do, that is, he wanted to have nonconsensual sex. In one sense this individualizes justice, for, if believed, the jury arguably knows something about the defendant that they did not know before. But what they know is ugly. It is arguably likely to anger the jury, to make the jury ignore defense evidence that the defendant is not like "typical" rapists or that, if he does share their qualities, he has other qualities that make him unlikely to think as they do, or he simply did not think that way on this occasion. Instead, he will be judged as a stereotype, a cipher, to be punished for membership in a group to which he apparently belongs, not for what he truly thought.

While this argument makes some sense, the empirical data and psychological theory suggest that expert character evidence often will not overwhelm the jury, although this question must ultimately be decided on a case-by-case basis. See supra text accompanying notes 616-640. It is precisely because a case-by-case judgment is appropriate that a flat bar would be unwise. Furthermore, under the proposal made here, prosecution-proffered expert character evidence must survive both Frye and the weighing of probative value against prejudice and catharsis-related values embodied in a broadened notion of "pragmatic relevancy." Moreover, pragmatic relevancy must be shown by clear and convincing evidence. Finally, an activist judge must have made careful efforts to
convinced” standard of proof for pragmatic relevancy therefore becomes pertinent only when the prosecution seeks to make first use of character evidence to prove mental state, because many courts find such evidence to be outside of Rule 404’s per se bar.665

On rebuttal, however, a lesser standard is justified. Once the defense has offered expert character evidence, the prosecution must be allowed to “repair the damage.”666 The door has been opened and fairness demands some opportunity for the prosecution to respond. More than this, if the prosecution cannot respond, a jury may receive incomplete information and may thereby be misled; facing only one real option, the jury is likely to choose it. The jury-centered value of full disclosure, serving both truth-seeking and political goals,667 is thus enhanced by the admission of prosecution rebuttal evidence. This justifies a reduced standard of proof.

But should the standard be so reduced as to place the prosecution in perfect parity with the defense? If not, a standard between preponderance and clear-and-convincing, perhaps labeled a “heavy preponderance,” would be appropriate. Because the cathartic value of individualized justice is particularly powerful for defense-proffered evidence—powerful enough to outweigh to a significant degree lingering concerns about reliability—this would be a wise compromise. The heavy preponderance standard would be somewhat higher than the defense standard, recognizing the reduced prosecution justification for such evidence and ensuring somewhat greater confidence in the evidence. On the other hand, the heavy preponderance standard would be significantly less than the clear-and-convincing standard for prosecution first-use, giving proper

reveal all evidentiary weaknesses to the jury, including cautions about ways they might abuse the evidence. This combination of safeguards seems ample to prevent the jury’s being misled by, or giving undue weight to, even prosecution-proffered group character evidence (for prosecution-proffered individual character evidence, there should by definition be even less reason to worry about stereotyping and de-individualizing justice). Indeed, absent extensive scientific support, it is unlikely that the “typical, violent rapist” example just given would survive that pragmatic relevancy weighing process.

In any event, because most courts find that Rule 404 does not flatly bar use of character evidence to prove mental state, see supra note 251 and accompanying text, it is likely that courts will be more receptive to arguments cast in terms of standards of proof than in terms of per se rules of exclusion.

665. See supra text accompanying notes 251-260.

666. Cf. McCord, Profiles, supra note 119, at 102 (making this point in discussing the relationship between reliability and necessity, not as to the standard of proof).

667. Political goals require giving the jury, as a political institution central to democratic decision making, more, not less, information to enable it to make decisions critical to a trial’s outcome whenever reasonably practical. See supra text accompanying notes 487-513.
IV. Conclusion

This Article has focused on the moral and social psychological need for individualized justice in criminal cases, a need recognized by the substantive rules of criminal law and which appropriately crafted rules of evidence can help to fulfill. Individualized justice requires consideration of a defendant's unique traits and circumstances to determine what that individual did or thought, not only what she should have done or thought. Yet our assembly-line justice system and common myths, prejudices, and misconceptions combine to lead to "normalized justice," to the trial of defendants for whom they are assumed to be and what they are assumed, therefore, to have done and thought. Such assumptions are largely based on what most people would do, clearly the opposite of individualized treatment.

This Article has thus focused on psychological character evidence—expert testimony as to the kind of person a defendant is—as a way to move judges and juries back toward an individualized assessment of the particular defendant before them. Psychological character testimony, because it focuses in part on a defendant’s history, on special traits and their consequences, and on the interaction between those traits and situational factors, offers the prospect of leading juries to think of defendants as special human beings, rather than as stereotypes. The clinical judgment often required in such testimony, however, can be of uncertain accuracy, part science but part art as well. This Article has thus sought to outline a general framework for revealing and testing the value of both the scientific and the artistic elements of such testimony.

Moreover, given the uncertainty about questions of accuracy, the Article has sought to explore justifications in addition to truth-seeking for admitting psychological character evidence. In particular, this Article focused on the "cathartic function"—the increased sense of individual and public respect for the justice system that comes from admitting evidence that enhances the prospect of individualizing justice.

The Article has advocated a judicial alternative to the "either/or" notion of admitting or excluding evidence, namely taking an activist role designed to improve the quantity and quality of the evidence presented to the jury. Strong political values are
served by giving the jury more, rather than less, information, wherever possible.

Finally, this Article has sought to contribute to a growing new understanding of the rules of evidence as not merely mechanical devices for reconstructing truth or as weapons of war for battle-starved litigators, but also as basic tools for enforcing society's moral and political values and for achieving that all-too-often-ignored goal in evidentiary discussions: justice.668

---

668. For an effective overview of the role of justice in legal decision making, including an analysis of some issues of individualized justice that, while not so identified by the authors, are fundamentally evidentiary questions, see Anthony D'Amato & Arthur J. Jacobson, Justice and the Legal System: A Coursebook (1992).