

Michigan Law Review

Volume 103 | Issue 4


2005

Psychology, Factfinding, and Entrapment

Kevin A. Smith

University of Michigan Law School

Follow this and additional works at: <https://repository.law.umich.edu/mlr>

 Part of the [Criminal Procedure Commons](#), [Law and Psychology Commons](#), and the [Law Enforcement and Corrections Commons](#)

Recommended Citation

Kevin A. Smith, *Psychology, Factfinding, and Entrapment*, 103 MICH. L. REV. 759 (2005).

Available at: <https://repository.law.umich.edu/mlr/vol103/iss4/3>

This Note is brought to you for free and open access by the Michigan Law Review at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Michigan Law Review by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

NOTE

Psychology, Factfinding, and Entrapment

Kevin A. Smith*

TABLE OF CONTENTS

INTRODUCTION	759
I. ATTRIBUTION THEORY & ENTRAPMENT	765
A. <i>Discounting the Situation and the Difficulty in Determining What Is "Inducement"</i>	767
B. <i>Inferring Disposition from Situational Behavior and the Difficulty in Determining Who Is "Predisposed"</i>	775
1. <i>Inferring Predisposition from Commission of the Offense</i>	775
2. <i>Failure to Apply Knowledge of Base Rates</i>	779
C. <i>Overconfidence and Reasonable Doubt</i>	781
D. <i>Channel Factors Applied to Sting Operations</i>	788
II. SOLUTIONS AND ALTERNATIVES	797
A. <i>An Objective Test Decided by a Jury</i>	797
1. <i>The Benefits of Jury Discussion</i>	798
2. <i>The Benefits of the Objective Test</i>	801
B. <i>Closing Arguments</i>	803
C. <i>Consideration of Inculcating and Exculpating Channel Factors in Motions for a Finding of Entrapment as a Matter of Law</i>	805
CONCLUSION	806

INTRODUCTION

Through the entrapment defense, the law acknowledges that criminal behavior is not always the result of a culpable mind, but is sometimes the result of an interaction between the individual and his environment. By limiting the amount of pressure and temptation that undercover agents may bring to bear on a target, the defense recognizes that the ordinary, law-abiding citizen can be persuaded, cajoled, or intimidated into criminal activity that, he would never

* Many thanks to Professor Phoebe Ellsworth for her comments on an earlier draft.

consider absent law-enforcement interference. Appropriate application of the defense requires, however, that courts be able to accurately separate the truly wicked from the merely weak-willed, and offensively coercive police conduct from that which merely convinces the criminal-minded to commit the crime *here and now* where he can more easily be caught. Two methods of making these distinctions have evolved: the subjective and objective tests.

In 1932, the Supreme Court first recognized the defense of entrapment in *Sorrells v. United States*,¹ creating what has since come to be known as the “subjective test.” The test has two elements: inducement and predisposition. The defendant must first provide evidence that the government induced him into committing the crime.² The inducement element focuses on the egregiousness of the pressure that the government brought to bear on the defendant. Although the degree of pressure required to establish inducement varies somewhat between courts, the element is generally satisfied where the police created a situation that posed a substantial risk that a hypothetical “ordinary law-abiding person” might be tempted to break the law.³ The Seventh Circuit has characterized the test for inducement as “whether the police employed tactics calculated to overcome the reluctance of a law-abiding citizen.”⁴ Similarly, the Eleventh Circuit describes inducement as government action that “created a substantial risk that the offense would be committed by a person other than one ready to commit it.”⁵ Because it emphasizes the reaction of an ordinary or reasonable citizen to police-created temptation, inducement is an objective inquiry.⁶

Once the defendant presents some evidence that the government did more than merely provide an opportunity to commit the crime — that is, offered some extra incentive for it — the burden shifts to the prosecution to rebut the defense by either proving beyond a

1. 287 U.S. 435 (1932).

2. *See, e.g.*, *United States v. Hanson*, 339 F.3d 983, 988 (D.C. Cir. 2003); *United States v. Ryan*, 289 F.3d 1339, 1343 (11th Cir. 2002).

3. *See, e.g.*, *United States v. Evans*, 924 F.2d 714, 717 (7th Cir. 1991) (describing inducement as “the sorts of promises that would blind the ordinary person to his legal duties”); *United States v. Andrews*, 765 F.2d 1491, 1499 (11th Cir. 1985) (holding that inducement occurs when the government created a substantial risk that the offense would be committed by someone “other than one ready to commit it”); *United States v. Kelly*, 748 F.2d 691, 697 (D.C. Cir. 1984) (“Inducement focuses on whether the government’s conduct could have caused an undisposed person to commit a crime.”); *United States v. Dickens*, 524 F.2d 441, 444 (5th Cir. 1975) (describing inducement as when “the Government’s conduct created a substantial risk that the offense would be committed by a person other than one ready to commit it”); PAUL MARCUS, *THE ENTRAPMENT DEFENSE* 60 (3d ed. 2002).

4. *Evans*, 924 F.2d at 717.

5. *Andrews*, 765 F.2d at 1499 (quoting *Dickens*, 524 F.2d at 444).

6. *Kelly*, 748 F.2d at 698.

reasonable doubt that the defendant was predisposed to commit the crime or that he was not induced.⁷ The predisposition element shifts the focus from how the police conduct would influence the ordinary person to whether this particular defendant would have committed an offense of the type charged in the absence of police inducement.⁸ If the defendant is predisposed to commit the crime, the defense fails. Accordingly, predisposition reduces to a question of “but-for” causation, looking to whether the defendant would have committed a crime but for the police pressure.⁹ Making this determination requires a prediction of how the defendant would have behaved had the police left him alone.¹⁰

The two elements thereby serve competing ends. The inducement inquiry, by requiring the prediction of an ordinary person’s behavior *with* police pressure, binds police tactics within a range that will not cause ordinary citizens to break the law. The predisposition inquiry, by allowing conviction notwithstanding the degree of inducement, prevents the would-be criminal from escaping liability for a crime that would likely have happened in any event. It therefore requires predicting *this* defendant’s behavior *without* police pressure.¹¹

7. See, e.g., *United States v. Hanson*, 339 F.3d 983, 988 (D.C. Cir. 2003); *United States v. Tom*, 330 F.3d 83, 89 (1st Cir. 2003).

8. MARCUS, *supra* note 3, at 63; Christopher D. Moore, Comment, *The Elusive Foundation of the Entrapment Defense*, 89 Nw. U. L. REV. 1151, 1164 (1995) (arguing that the Court’s current position on predisposition asks “whether the defendant would have committed the crime absent government involvement”).

9. See *Sorrells v. United States*, 287 U.S. 435, 444-45 (1932) (stating that entrapment prevents the punishment of a defendant “for the commission of an offense of the like of which he had never been guilty, either in thought or in deed, and evidently never would have been guilty of if the officers of the law had not inspired . . . him to attempt to commit it”); *United States v. Manzella*, 791 F.2d 1263, 1269 (7th Cir. 1986) (“As a defense to a criminal prosecution ‘entrapment’ means the government’s inducing a person to commit a crime who was not predisposed to commit it — in other words, who would not have committed it but for the particular inducement that the government held out.”); MARCUS, *supra* note 3, at 63; John D. Lombardo, Comment, *Causation and “Objective” Entrapment: Toward a Culpability-Centered Approach*, 43 UCLA L. REV. 209, 235 (1995).

10. See *United States v. Hollingsworth*, 27 F.3d 1196, 1201-02 (7th Cir. 1994) (en banc). In *Hollingsworth*, the court found entrapment as a matter of law where, despite his willingness when given the opportunity to launder money, the defendant lacked predisposition to commit the charged offense because, without the government’s assistance, he never would have had the opportunity to do so. *Id.* at 1202 (“[The defendants] had no prayer of becoming launderers without the government’s aid.”).

11. See Phillip Mullock, *The Logic of Entrapment*, 46 U. PITT. L. REV. 739, 745 (1985) (“[T]o say that [the defendant] is disposed [to violate the statute] is to say that *if* a suitable opportunity presents itself, [the defendant] will probably [violate the statute].”). Others have erroneously claimed that a defendant is predisposed so long as under some circumstance he was willing to commit the crime. See Ronald J. Allen et al., *Clarifying Entrapment*, 89 J. CRIM. L. & CRIMINOLOGY 407, 413 (1999) (arguing that predisposition, as applied by some courts, is “an existential fallacy” because everyone has a price at which they will commit a crime); Jonathan C. Carlson, *The Act Requirement and the Foundations of the Entrapment Defense*, 73 VA. L. REV. 1011, 1040 (1987) (making the same argument). This reading of predisposition is circular. The jury knows the defendant has a price because they know he

Many states have rejected the subjective test and adopted what has become known as the "objective test."¹² The objective test forgoes any inquiry into the defendant's predisposition, and merely asks whether the police conduct posed a substantial risk that the crime would be committed by someone otherwise not ready to do so.¹³ This focus on police conduct is essentially the inducement inquiry of the subjective test.¹⁴ Some states also require that the defendant prove a causal connection between the improper police conduct and the commission of the crime.¹⁵ While the subjective test almost always leaves the question of entrapment to the jury, in most states employing the objective test, the question is exclusively for the judge.¹⁶

Although accurate statistical evidence of the frequency with which the entrapment defense succeeds is difficult to come by, anecdotal evidence suggests that it is seldom successful. One survey of practicing criminal defense attorneys described it as "judicially unpopular,"¹⁷ best used only "in desperate circumstances,"¹⁸ or "in a few cases with ideal

committed the offense. The relevant question is what is the defendant's price in relation to the unmanipulated world? Because entrapment is concerned with preventing the police from manufacturing crime that would not otherwise occur, the best reading of the doctrine is that the predisposed is one who would have committed the crime under the ordinary circumstances of his daily life. The predisposed is, therefore, not one who would commit the crime in exchange for some extraordinarily high price, but one who would do so when provided with the incentives available in his ordinary environment.

12. Because the entrapment defense is based on a presumed congressional intent that criminal statutes were not meant to reach acts instigated by the government, state courts are free to adopt other versions of the defense in interpreting their own state law. *See* *Sherman v. United States*, 356 U.S. 369, 379-82 (1958) (Frankfurter, J., concurring); *Sorrells v. United States*, 287 U.S. 435, 446-51 (1932).

13. *State v. Agrabante*, 830 P.2d 492, 499 (Haw. 1992); *Grossman v. State*, 457 P.2d 226, 229 (Alaska 1969); MODEL PENAL CODE § 2.13 (1985).

14. Compare MODEL PENAL CODE § 2.13 (1985) (describing the objective test for entrapment as when a police officer induces the defendant by "employing methods of persuasion or inducement that create a substantial risk that [the] offense will be committed by persons other than those who are ready to commit it"), with *United States v. Andrews*, 765 F.2d 1491, 1499 (11th Cir. 1985) (holding that inducement occurs under the subjective test when the government created a substantial risk that the offense would be committed by someone "other than one ready to commit it"), and sources cited *supra* note 3.

15. *See, e.g., Municipality of Anchorage v. Flanagan*, 649 P.2d 957, 961 (Alaska Ct. App. 1982) (holding that the objective defense of entrapment was not available where the inducement did not cause the defendant's crime); *People v. Barker*, 293 N.W.2d 787, 788 (Mich. Ct. App. 1980) ("Defendant's selling of stolen property occurred *after* the breaking and entering for which he was charged. Thus, it cannot rationally be argued that the subsequent police conduct caused defendant's prior offense. Defendant was not charged with any offenses arising out of the police sale, therefore, there was no possible entrapment."), *rev'd on other grounds*, 306 N.W.2d 100 (Mich. 1981); *see also* Lombardo, *supra* note 9, at 238-40.

16. MARCUS, *supra* note 3, at 247; *e.g., State v. Valdez-Molina*, 897 P.2d 993, 995 (Idaho 1995); *People v. D'Angelo*, 257 N.W.2d 655 (Mich. 1977); *Grossman v. State*, 457 P.2d 226, 230 (Alaska 1969).

17. Ben A. Hardy, *The Traps of Entrapment*, 3 AM. J. CRIM. L. 165, 189 (1974).

18. *Id.* at 165.

facts,” where “plea bargaining has proved unsuccessful,”¹⁹ and “no other defense is possible.”²⁰ Another survey of State’s Attorneys in Chicago reveals a perception that in narcotics sales cases where the defendant pleads entrapment, “the jury will convict almost every time.”²¹ The author of one police manual on the execution of sting operations states that he has “never, in hundreds of cases, ever lost one to entrapment,” and that in all the sting operations he has studied, he has “not heard of a single case being lost to a defense of entrapment.”²²

Absent information concerning how frequently the entrapment defense *should* be successful, statistics on how often it *is* successful provide little guidance. If we are confident that the police are not placing undue pressure on the targets of sting operations, then the defense should seldom succeed. But one study of police training manuals finds that they contain no “significant discussions of [entrapment].”²³ The discussion that does exist is lacking in meaningful advice to the practicing police officer.²⁴ This dearth of meaningful training on entrapment, the authors conclude, is because the concept is not a significant limitation on police practices.²⁵

Two explanations for the defense’s ineffectiveness have been widely offered. First, for entrapment to be credibly argued, the criminal act must be admitted.²⁶ While a defendant can, in theory, both deny committing the criminal act and plead entrapment, such a strategy destroys the defendant’s credibility.²⁷ Even when the police admit actions that are plausibly inducing, a defendant may elect to forgo pleading entrapment and instead deny committing the offense. Second, the subjective form of the defense invites the prosecution to introduce a variety of evidence relevant to the defendant’s predisposition such as his criminal record; prior bad acts, including uncharged crimes; and reputation testimony, which would otherwise

19. *Id.* at 165.

20. *Id.* at 165, 189.

21. David P. Bancroft, Comment, *Administration of the Affirmative Trap and Doctrine of Entrapment: Device and Defense*, 31 U. CHI. L. REV. 137, 162 (1963); see also Roger Park, *The Entrapment Controversy*, 60 MINN. L. REV. 163, 267 n.339 (1976) (describing an informal survey of Minnesota criminal attorneys and their experience with the entrapment defense).

22. STEVEN K. FRAZIER, *THE STING BOOK* 137 (1994).

23. LAWRENCE P. TIFFANY ET AL., *DETECTION OF CRIME* 211 (1976).

24. *Id.*

25. *Id.* Members of a vice squad of the Chicago Police Department informed the authors that entrapment was “not treated formally or informally as a problem of concern.” *Id.* at 211 n.18.

26. See *Mathews v. United States*, 485 U.S. 58, 67 (1988) (Scalia, J., concurring).

27. *Id.*

be excluded as propensity evidence.²⁸ The introduction of this evidence is often criticized because it invites the jury to convict based on the defendant's bad character or uncharged bad acts, even if they are not convinced that he was predisposed to commit the offense.²⁹

This Note argues that there is another reason entrapment claims seldom succeed: the entrapment defense requires that factfinders make causal and predictive determinations that human beings are unable to make with accuracy, and the inherent inaccuracy works to the defendant's disadvantage. The entrapment defense suffers from what has been called "the fundamental attribution error" — a pervasive bias to see people as causal agents in their environment.³⁰ Studies in social psychology demonstrate that people consistently overestimate the role of dispositions and underestimate the role of the situation when making predictions about what a particular defendant or an ordinary person would do in a given situation — the sorts of inquiries that an entrapment defense requires. Compounding the problem, people consistently overestimate the accuracy with which they can make these sorts of predictions, thereby depriving the defendant of the benefit of the reasonable doubt. The effect of these inaccuracies is to consistently impair the efficacy of the defense.³¹

This Note relies on studies in social psychology to argue that the ineffectiveness of the entrapment defense can be explained by commonly held cognitive biases that sway the jury toward conviction. Part I argues that the fundamental attribution error biases the factfinding in entrapment cases against the defendant. Studies in social psychology reveal that factfinders do a consistently poor job of determining what types of situations are inducing, that they are prone to infer dispositions from manifestly situational behavior, and are consistently overconfident in their ability to make such findings accurately. Part II proposes several methods for mitigating the effects

28. Compare FED. R. EVID. 404(a) ("Evidence 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 . . ."), with *Sorrells v. United States*, 287 U.S. 435, 451 (1932) ("[I]f the defendant seeks acquittal by reason of entrapment he cannot complain of an appropriate and searching inquiry into his own conduct and predisposition as bearing upon that issue."). See generally W.H. Johnson, III, Note, *Proving a Criminal Predisposition: Separating the Unwary Innocent from the Unwary Criminal*, 43 DUKE L.J. 384 (1993).

29. *Sherman v. United States*, 356 U.S. 369, 383 (1958) (Franklin, J., concurring).

30. See generally LEE ROSS & RICHARD E. NISBETT, *THE PERSON AND THE SITUATION: PERSPECTIVES OF SOCIAL PSYCHOLOGY* (1991) (discussing the fundamental attribution error).

31. For an interesting argument that the fundamental attribution error has produced a systemic tendency within the substantive criminal law to find criminal culpability when harm occurs, see Donald A. Dripps, *Fundamental Retribution Error: Criminal Justice and the Social Psychology of Blame*, 56 VAND. L. REV. 1383 (2003). This Note, however, takes a less novel approach, examining solely the fundamental attribution error's effect on factfinding, and not how it has shaped legal doctrine.

of the fundamental attribution error: improving jury instructions, adopting the objective test for entrapment as a jury question, and increasing police and judicial sensitivity toward the use of channel factors³² in sting operations.

I. ATTRIBUTION THEORY & ENTRAPMENT

Implicit in both the entrapment defense and classic attribution theory is the understanding that behavior is produced by a combination of an individual's personal characteristics (his disposition) and his environment (the situation). All of an individual's actions fall along a causal spectrum. On one end lie those actions that are completely situational — under the circumstances, anybody would have behaved the same way.³³ At the other lie those actions that, under the circumstances, differ from those that most people would take under the same circumstances. We attribute this deviation from the norm to the unique disposition of the actor.³⁴ Relative to the ordinary man, the actor had a "preference for" or a "disposition toward" the particular act.

Determining where on this causal spectrum any one person's particular action lies is a highly inaccurate task. Social psychology teaches that the attribution of causality between situation and disposition is biased by the fundamental attribution error, which results in causal attributions that lean excessively toward the dispositional end of the spectrum.³⁵ In other words, "in a social setting in which either a person or some situational variable is a plausible causal candidate for an outcome, there exists a general bias to see people as causal agents, particularly their enduring dispositional

32. A channel factor is a very small situational difference that produces surprisingly large changes in behavior. See discussion *infra* Part I, Section D.

33. See Harold H. Kelley, *Attribution Theory in Social Psychology*, in 15 NEBRASKA SYMPOSIUM ON MOTIVATION, 192, 208-09 (David Levine ed., 1967) ("[T]he action yields information about the actor's idiosyncratic intentions only insofar as the effects are not those that people in general would have produced under similar circumstances."); Richard E. Nisbett & Eugene Borgida, *Attribution and the Psychology of Prediction*, 32 J. PERSONALITY & SOC. PSYCHOL. 932, 932 (1975) ("[T]he stimulus rather than the actor will be seen as the chief cause . . . when most people respond to the entity in the same way the actor does.").

34. See Edward E. Jones & Victor A. Harris, *The Attribution of Attitudes*, 3 J. EXPERIMENTAL SOC. PSYCHOL. 1 (1967).

35. See generally RICHARD E. NISBETT & LEE ROSS, HUMAN INFERENCE: STRATEGIES AND SHORTCOMINGS OF SOCIAL JUDGMENT (1980); Lee Ross, *The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process*, in 10 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY 173 (Leonard Berkowitz ed., 1977); Edward E. Jones & Richard E. Nisbett, *The Actor and the Observer: Divergent Perceptions of the Causes of Behavior*, in ATTRIBUTION: PERCEIVING THE CAUSES OF BEHAVIOR 79 (Edward E. Jones et al. eds., 1971); Edward E. Jones, *The Rocky Road from Acts to Dispositions*, 34 AM. PSYCHOLOGIST 107 (1979).

attributes.”³⁶ This is not to say that disposition is not a determinant of behavior, but merely that it is often less so than is commonly believed. The fundamental attribution error has two components: First, people have an inflated belief in the importance of individual character differences and dispositions.³⁷ Second, they underestimate the degree to which situational factors influence behavior.³⁸

The fundamental attribution error poses a particular threat to the entrapment defense because the defense requires the factfinder to separate the situational and dispositional causes of behavior. Section I.A argues that factfinders have a difficult time determining when an ordinary person would be tempted by a sting operation. Determining whether a situation is inducing requires predicting how the ordinary person would respond to the police sting. Numerous studies suggest, however, that people consistently underestimate how easily behavior can be manipulated by situational factors.

Section I.B argues that a factfinder may have similar difficulty in determining whether a defendant is predisposed to a particular crime. This element of the defense requires the factfinder to consider a variety of character evidence and determine whether the defendant was the sort of person who was likely to commit the offense anyway. This is akin to predicting the likelihood that the defendant would commit the crime even if he had never wandered into the police sting operation. Studies suggest, however, that even when given reliable character evidence, people do a very poor job of predicting how others will behave.

Section I.C contends that the fundamental attribution error produces overly confident factual findings in entrapment cases. Factfinders are more confident in their ability to identify dispositions and predict how the ordinary person will respond to a particular situation than their performance warrants. This overconfidence works to the defendant's disadvantage by depriving him of the benefit of the reasonable doubt.

Finally, Section I.D discusses the use of channel factors in sting operations. Social psychology provides a variety of methods of designing sting operations that have the tendency to produce drastic changes in behavior through subtle situational manipulation. Those manipulations that are particularly difficult for factfinders to account

36. Shelley E. Taylor, *The Availability Bias in Social Perception and Interaction, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* 190, 192-93 (Daniel Kahneman et al. eds., 1982). For further discussion and an historical background of the fundamental attribution error, see Daniel T. Gilbert & Patrick S. Malone, *The Correspondence Bias*, 117 *PSYCHOL. BULL.* 21 (1995).

37. LEE ROSS & RICHARD E. NISBETT, *THE PERSON AND THE SITUATION: PERSPECTIVES OF SOCIAL PSYCHOLOGY* 4 (1991).

38. *Id.*

for are “channel factors,” and pose a particular threat to the entrapment defense.

A. *Discounting the Situation and the Difficulty in Determining What Is “Inducement”*

When the factfinder is required to determine whether the defendant was induced into committing a crime, it must decide whether the police created a situation that would potentially cause a law-abiding citizen to break the law.³⁹ This exercise is really a form of prediction: the factfinder imagines a hypothetical reasonable person and predicts how he would respond to a given police inducement. Studies in social psychology demonstrate, however, that subjects often highly underestimate the degree to which the ordinary person’s behavior can be altered by subtle situational manipulations.⁴⁰ This substantially decreases the likelihood that a judge or jury will find that police pressure was sufficient to cause a reasonable person to deviate from his ordinary law-abiding ways. Instead, because it has excessively discounted situational factors as causes for the crime, the factfinder is overly likely to attribute the crime in question to a criminal disposition possessed by the defendant.

In a groundbreaking study, John Darley and Daniel Batson explored the degree to which personal disposition and situational factors determine behavior through an experimental-recreation of the parable of the Good Samaritan.⁴¹ The subjects of the experiment were Princeton University theological seminary students.⁴² In the first stage, each subject was given a questionnaire to complete regarding the reasons for their decision to attend the seminary.⁴³ Specifically, the

39. See sources cited *supra* note 3 and accompanying text.

40. See, e.g., Walter Mischel & Philip K. Peake, *Beyond Déjà Vu in the Search for Cross-Situational Consistency*, 89 PSYCHOL. REV. 730, 730 (1982) (“[C]ompelling intuitive evidence supports the enduring conviction that people are characterized by broad dispositions revealed in extensive cross-situational consistency.”).

41. John M. Darley & C. Daniel Batson, *From Jerusalem to Jericho: A Study of Situational and Dispositional Variables in Helping Behavior*, 27 J. PERSONALITY & SOC. PSYCHOL. 100 (1973). The parable of the Good Samaritan is as follows:

“And who is my neighbor?” Jesus replied, “A man was going down from Jerusalem to Jericho, and he fell among robbers, who stripped him and beat him, and departed, leaving him half dead. Now by chance a priest was going down the road; and when he saw him he passed by on the other side. So likewise a Levite, when he came to the place and saw him, passed by on the other side. But a Samaritan, as he journeyed, came to where he was; and when he saw him, he had compassion, and went to him and bound his wounds . . . Which of these three, do you think, proved neighbor to him who fell among the robbers?”

Luke 10:29-36 (RSV).

42. Darley & Batson, *supra* note 41, at 102.

43. *Id.*

researchers requested that the subject identify a dispositional trait that best explained the motivation for his religious training.⁴⁴ In Stage II, each subject was individually instructed to prepare a short talk to be presented in a nearby building.⁴⁵ After being given directions to the building, one-third of the subjects were told, ““Oh, you’re late. They were expecting you a few minutes ago. . . . [Y]ou’d better hurry;”” another third were told, ““[They’re] ready for you, so please go right over;”” and the last third were told, ““It’ll be a few minutes before they’re ready for you, but you might as well head on over.””⁴⁶ On the way, while passing through an alley, the participants encountered a man slumped in a doorway, coughing and groaning.⁴⁷ The subjects were given scores according to how much assistance, if any, they offered to the apparently sick man.⁴⁸

Darley and Batson discovered that they could manipulate how most seminarians reacted to the sick man by varying the degree of time-pressure they were under. While 63 percent of early seminarians, and 45 percent of the on-time seminarians stopped to help, only 10 percent of the late seminarians were helpful.⁴⁹ The dispositional variable — the nature of the participant’s religious orientation — had no statistically significant effect.⁵⁰ Because most seminarians did not help when they were hurried, but few did when hurried, Darley and Batson were able to “induce” either helpful or unhelpful behavior by altering the situational manipulation.

Observers of the Good Samaritan experiment were not, however, able to easily perceive the true causes of the seminarians’ behavior. Expanding on the Good Samaritan study, Paula Pietromonaco and Richard Nisbett provided half of their subjects with a copy of Darley and Batson’s results to read.⁵¹ The researchers then asked all of the subjects to make predictions about helping behavior in two situations.⁵² One situation, very similar to the Good Samaritan study,

44. *Id.* at 103. The three possible motivations for their training were: a) a means to another end such as salvation; b) religion as an end in itself; or c) religion as a quest for meaning in the world. *Id.*

45. *Id.*

46. *Id.* at 103-04.

47. *Id.*

48. *Id.*

49. *Id.* at 105.

50. *Id.* at 106 (“No correlation between the various measures of religiosity and [the occurrence of helping behavior] ever came near statistical significance . . .”).

51. Paula R. Pietromonaco & Richard E. Nisbett, *Swimming Upstream Against the Fundamental Attribution Error: Subjects’ Weak Generalizations from the Darley and Batson Study*, 10 SOC. BEHAV. & PERSONALITY 1, 2 (1982).

52. *Id.*

involved a woman with an injury requesting assistance.⁵³ The second was a pregnant woman whose car had broken down.⁵⁴ Each subject then estimated the likelihood that the average member of the target population would help.⁵⁵ Before answering, those who had received the Good Samaritan study were asked to recall and consider it.⁵⁶

The subjects demonstrated a poor grasp of what causes a person to help another. For those subjects who had not read the Good Samaritan study, the estimates regarding which seminarians would help were virtually the opposite of reality. Uninformed subjects believed that being in a hurry would have virtually no effect on the decision, and predicted that about 80 percent of seminarians in both the hurried and unhurried groups would stop to help.⁵⁷ The uninformed subjects also erroneously believed that religious disposition would have a pronounced effect, producing nearly an eighteen-percentage point difference in the proportion of seminarians offering help.⁵⁸ The results suggest that the observers of this experiment are inclined to attribute helping behavior to dispositional qualities of the actors and not to salient situational factors. Although the primary determinant of helping behavior was whether or not the experimenters “induced” the seminarian into not helping by hurrying him, observers believed that this factor would be irrelevant. Instead, they placed great emphasis on what turned out to be the statistically insignificant religious disposition of the seminarian.

This tendency is especially resilient to remedial efforts. Even those subjects who had read and been instructed to recall the Good Samaritan study fared poorly. They were no less likely to use religious motivation as a factor in predicting when seminarians would help than the uninformed subjects.⁵⁹ In addition, while Darley and Batson had demonstrated a fifty-three percentage point difference in helping rates between hurried and unhurried groups, the estimates of the informed subjects regarding which seminarians would help only showed a nineteen point difference.⁶⁰ The researchers concluded that “the failure of subjects to generalize even to a highly similar situation,

53. *Id.*

54. *Id.*

55. For example, subjects were asked, “What percentage of the seminary students would help?” or “What percentage of New Jersey males would help?” *Id.*

56. *Id.*

57. *Id.* at 3 tbl.1.

58. *Id.* at 3.

59. *Id.* (“No interaction effects were obtained between the effect of being informed and personality type or for any interactions including these two variables. Thus, informed subjects continued to use the personality variable for predictions and were not significantly less inclined to do so than uninformed subjects.”).

60. *Id.* at 3 tbl. 1.

despite the presence of what was surely a very strong experimental demand to do so, suggests that the fundamental attribution error may be quite resistant to the data and arguments of the social scientist.”⁶¹

Like the subjects of these studies, the factfinder in an entrapment case is presented with situational and dispositional data and required to determine the cause of a particular behavior. Factfinders must first determine how a reasonable person would respond to the police sting operation. Second, given the evidence concerning the defendant’s character and disposition, the factfinder must decide whether the crime would have occurred even absent the inducement. What these studies suggest for the entrapment defense is that judges and juries may perform quite poorly at these tasks. The drastic behavioral effects produced by subtle situational manipulations may go unnoticed. Instead, factfinders may overestimate the probative value of the more salient propensity evidence and explain the defendant’s behavior as a product of a criminal disposition. These studies suggest, however, that dispositional evidence is a much less valuable means of determining the cause of behavior than the average judge or jury is likely to realize.

The results of Darley & Batson’s study might plausibly be explained by overconfidence in the accuracy of the seminarian’s self-reported, dispositional evidence. In an entrapment case, the defendant’s disposition will likely be proven in part through the testimony of character witnesses.⁶² One might plausibly hope that the opinions of others would provide a more accurate assessment of an actor’s disposition than potentially self-serving self-reports, thereby improving the accuracy of predictions relying on dispositional evidence. The fundamental attribution error has, however, been demonstrated to be a robust phenomenon even when dispositional evidence is derived from a more objective source.

Newton, Griffin and Ross gave two groups of subjects the chance to donate to a food drive.⁶³ Subjects were ranked by their peers according to the likelihood that they would donate to a food drive.⁶⁴ Those voted “least likely” were placed in one group; those voted “most likely” in another.⁶⁵ Half of the subjects in each group were sent both a letter personally addressed to them requesting that they donate and a map showing the location of the collection box, and they

61. *Id.* at 4.

62. Thomas J. Reed, *The Character Evidence Defense: Acquittal Based on Good Character*, 45 CLEV. ST. L. REV. 345, 392-93 (1997).

63. ROSS & NISBETT, *supra* note 30, at 132-33 (describing E. Newton et al., *Actual Versus Estimated Impact of Person and Situation in Determining Pro-Social Behavior* (1988) (unpublished Manuscript, on file with Stanford University)).

64. *Id.* at 132.

65. *Id.*

received follow-up phone call reminding them to donate.⁶⁶ The other half of each group was merely given a form letter addressed to “Dear Student,” requesting that they donate.⁶⁷

The peers who had nominated each subject to his “least likely” or “most likely” group were then asked to predict whether the subject would donate, given the situational group he was placed in.⁶⁸ The nominators strongly believed that the disposition of the subject would be predictive of donation, but that the situational group he was placed in would be largely irrelevant.⁶⁹ Specifically, they estimated that a vast majority of the “most likely” contributors, but only a small minority of the “least likely” contributors, would donate.⁷⁰ These predictions did not substantially vary depending on whether the subject would receive a map, phone call, and personalized letter, or merely a form letter.⁷¹

It was, however, the situation that proved a stronger determinant of behavior than a person’s perceived disposition to donate. None of the “least likely” subjects and only 8 percent of the “most likelies” who received only the form letter donated to the drive,⁷² but 25 percent of the “least likelies” and 42 percent of the “most likelies” donated when they received the map, personalized letter, and follow-up phone call.⁷³ While the study revealed that both disposition and situation are important factors in predicting a person’s behavior, the situational variable, at least in this instance, proved more important than any character trait salient to his peers.⁷⁴

Even when given character evidence provided by the actor’s peers, subjects in the donation study performed poorly at determining when a peer would be “induced” into donating and when he would do so

66. *Id.*

67. *Id.*

68. *Id.* at 133.

69. *Id.*

70. The nominators predicted likelihoods of 83% for those who received the phone call, personalized letter, and map, and an 80% likelihood for those that received only the form letter. *Id.* The subjects predicted that the “least likely” contributors would have a 17% likelihood of donating when they received the map, personalized letter, and follow-up phone call, and a 16% likelihood of donating where they only received the form letter. *Id.*

71. *See id.*

72. *Id.*

73. *Id.*

74. Ross and Nisbett concluded:

The evidence, in fact, is consistent with an extreme version of the fundamental attribution error. People readily make trait ascriptions from data that permit only a situational interpretation or, at most, the interpretation that that the actor behaves in a particular way in a particular type of situation. These trait ascriptions are then used as the basis for yet further predictions, which, again are characterized by little attention to situational factors.

Id.

due to his own disposition. Predictors were able to somewhat separate those of their peers who had donating disposition from those who did not, but they overestimated the role that disposition would have in their peer's decision. Similarly, predictors weakly perceived that receiving a personalized letter, phone call, and map were situational variables that would have some effect on the likelihood of donation, but they vastly underestimated their role in the experiment's outcome.

Extrapolating from studies of helping behavior to the realm of crime prevention is, of course, fraught with uncertainty. While it may be fairly easy to induce an ordinary person into donating to a can drive, producing the sort of antisocial theft and drug offenses typically targeted by sting operations may not be as easy. It may seem that criminal activity is so beyond the realm of an ordinary person's typical behavior that the sort of subtle situational manipulations that had such drastic effects in these studies would be ineffectual at instigating crime. Instead, the more that crime deviates from the person's ordinary behavior, the more drastic and obvious the inducement might have to be.

As intuitive as this argument may be, one of the most famous studies in social psychology, Stanley Milgram's *Behavioral Study of Obedience*, refutes it.⁷⁵ Milgram's subjects were instructed that they were participating in a study of memory and learning.⁷⁶ They were partnered with one of Milgram's accomplices, masquerading as another subject.⁷⁷ In what appeared to be a random manner, subjects were assigned to the role of teacher, and accomplices to the role of learner.⁷⁸ Accomplice and subject were taken to a room where the learner was strapped into an electric chair.⁷⁹ The teacher was then taken to the shock generator in an adjoining room connected via an intercom.⁸⁰ The teacher was instructed to ask the learner a series of memory questions.⁸¹ Each time the learner answered one incorrectly, the teacher was to press a button to administer a shock and then turn a dial to increase the voltage for the next wrong answer.⁸² Once the

75. Stanley Milgram, *Behavioral Study of Obedience*, 67 J. ABNORMAL & SOC. PSYCHOL. 371 (1963) [hereinafter Milgram, *Behavioral Study*]; see also STANLEY MILGRAM, OBEDIENCE TO AUTHORITY (1974) [hereinafter MILGRAM, OBEDIENCE TO AUTHORITY]; Stanley Milgram, *Some Conditions of Obedience and Disobedience to Authority*, 18 HUM. REL. 57 (1965) [hereinafter Milgram, *Some Conditions*].

76. Milgram, *Behavioral Study*, *supra* note 75, at 372.

77. *Id.* at 373.

78. *Id.*

79. *Id.* At this point, the teacher was given a 45 volt sample shock to convince him of the authenticity of the experiment. *Id.*

80. *Id.*

81. *Id.*

82. *Id.* at 373-74.

shock reached the level of 300 volts, the learner began to bang on the adjoining wall.⁸³ The pounding was repeated after a 315 volt shock was administered, and the learner ceased to respond either to questions or further shocks after that.⁸⁴ If the teacher continued asking questions, the learner ceased to respond, and the teacher was instructed to treat no answer as a wrong answer.⁸⁵ The experiment ended either when the teacher refused to continue or when he administered the maximum shock of 450 volts, two steps beyond the level labeled "Danger: Severe Shock."⁸⁶

Milgram sought to discover what portion of teachers would call a halt to the experiment and refuse to continue administering shocks to a non-responsive learner. The results of the study were startling. Only one in eight subjects refused to continue administering shocks after the learner began banging on the wall,⁸⁷ and in total, 35 percent at some time refused to continue shocks.⁸⁸ But 65 percent never refused to comply, administering ten more shocks of increasing intensity after the learner began banging on the wall and eight more shocks after he stopped responding at all.⁸⁹ With few exceptions, the teachers believed the shocks they administered were real and painful.⁹⁰ Many fully obedient teachers orally protested, but continued with the experiment when the researcher accepted responsibility and instructed them to continue.⁹¹

Anticipating that the results might be counterintuitive, Milgram measured the difference between commonly expected obedience and actual obedience. Prior to conducting the experiment, Milgram

83. *Id.* at 374.

84. *Id.*

85. *Id.*

86. *See id.* at 372, 376.

87. *Id.* at 375.

88. *Id.*

89. *Id.* The maximum shock that these subjects administered was two notches beyond where the dial was labeled "Danger: Severe Shock." *Id.* at 376.

90. *Id.* at 375.

91. *See* MILGRAM, OBEDIENCE TO AUTHORITY, *supra* note 75, at 153-68; Milgram, *Some Conditions*, *supra* note 75, at 67. Milgram repeated the study several more times under different conditions. *See* MILGRAM, OBEDIENCE TO AUTHORITY, *supra* note 75; Milgram, *Some Conditions*, *supra* note 75. In one variation, Milgram placed the learner in the same room as the teacher. Milgram, *Some Conditions*, *supra* note 75, at 62. In this version, the learner would only receive a shock if his hand was pressed against a metal plate on the arm of the electric chair. *Id.* At the 150-volt level, the learner demanded to be let free and refused to place his hand on the shock-plate. *Id.* The experimenter ordered the teacher to force the learner's hand onto the plate and administer the remaining shocks. At 300 volts the learner refused to answer any more questions and insisted he was no longer a willing participant. *Id.* at 60. Even under these conditions, 30 percent of the teachers completed the entire experiment. *Id.* at 62.

distributed a description of the proposed study and a questionnaire to fourteen Yale senior psychology majors.⁹² They were asked to reflect on the study and predict the behavior of 100 hypothetical teachers.⁹³ All fourteen senior psychology students agreed that only a small minority of teachers would continue to the maximum level of shock.⁹⁴ The estimates ranged only from 0 percent to 3 percent.⁹⁵ Milgram posed the same question to his colleagues, and “the most general feeling was that few if any subjects would go beyond the designation Very Strong Shock.”⁹⁶ The prediction of forty psychiatrists surveyed by Milgram was that only 3.73 percent of teachers would continue after the learner began pounding on the wall, and a mere one-tenth of 1 percent of subjects would administer the maximum level of shock.⁹⁷

Milgram’s experiment suggests that most ordinary, reasonable people can be induced into committing acts that, had circumstances been as the subjects believed, would be criminal. Moreover, Milgram was able to induce this behavior through situational manipulations that few professional psychologists and psychology students — presumably a group well-equipped to identify inducement — could accurately evaluate. Clearly, however, Milgram’s teachers were induced into shocking the learner. The experiment went beyond posing a substantial risk that an ordinary person would administer the series of shocks; it produced a situation in which nearly two-thirds did! Assuming for the moment that Milgram was a government agent and the teacher was charged with an offense related to the shocks, no jury could rationally conclude beyond a reasonable doubt that there was no inducement. Yet the predictions of Milgram’s colleagues and students suggest that hardly any judge or jury would find that such a defendant was induced. Instead, most courts would likely hold that the experimental design failed to do more than merely provide an “opportunity” to commit the offense and is therefore insufficient to even raise the issue of entrapment.⁹⁸

92. Milgram, *Behavioral Study*, *supra* note 75, at 375.

93. *Id.*

94. *Id.*

95. *Id.*

96. *Id.*

97. Milgram, *Some Conditions*, *supra* note 75, at 72-73 & fig.3.

98. See *Sorrells v. United States*, 287 U.S. 435, 441 (1932) (“It is well settled that the fact that officers or employees of the government merely afford opportunities or facilities for the commission of the offense does not defeat the prosecution.”).

B. *Inferring Disposition from Situational Behavior and the Difficulty in Determining Who Is “Predisposed”*

The flipside of the propensity to underestimate the role that situational variables play in determining behavior is that when behavior is observed, the observer is too likely to infer that the actor possesses a corresponding character trait.⁹⁹ When people underestimate the power of situations, they will be prone to make unwarranted dispositional inferences about actors who violate the erroneous expectations that such underestimates create.¹⁰⁰ Observers also fail to properly utilize knowledge of base rates — the proportion of other people who behaved similarly — when making their dispositional attributions.¹⁰¹ Observers are not substantially less likely to infer character traits from a person’s behavior when they know that most or all other people behaved similarly in the same situation.¹⁰²

This phenomenon may undermine the factfinder’s ability to accurately determine the predisposition of the defendant in an entrapment case. First, a factfinder may be excessively prone to infer that the entrapped defendant was predisposed to commit the crime based on his induced commission of the offense. Second, the factfinder is likely to undervalue evidence that, given the degree of inducement, most other people would also have committed the offense.

1. *Inferring Predisposition from Commission of the Offense*

Studies in social psychology and about the fundamental attribution error confirm a long-suspected pitfall of the entrapment defense: factfinders may infer that the defendant was predisposed to commit the offense merely from the fact that he did so in response to the police inducement.¹⁰³ The fact that he committed the crime is often the

99. Jones, *supra* note 35, at 113-16 (discussing the overattribution effect).

100. See generally Melvin Snyder & Edward E. Jones, *Attitude Attribution When Behavior Is Constrained*, 10 J. EXPERIMENTAL SOC. PSYCHOL. 585 (1974) (providing evidence that people tend to make dispositional attributions to explain behavior, underestimating the role of environmental constraints).

101. See generally Nisbett & Borgida, *supra* note 33 (arguing that subjects do not make use of base-rate information in making predictions).

102. *Id.*

103. See Bennett L. Gershman, *Abscam, The Judiciary, and the Ethics of Entrapment*, 91 YALE L.J. 1565, 1581 (1982); Hardy, *supra* note 17 at 187 (“[O]ne of the most common methods of proving a defendant’s predisposition is by showing his ready complaisance to commit the crime or, in effect, the speed and ease with which he complies with an agent’s or informant’s request.”) Stephen A. Miller, Comment, *The Case for Preserving the Outrageous Government Conduct Defense*, 91 NW. U. L. REV. 305, 329-30 (1996) (“Juries may infer predisposition simply from a defendant’s acceptance of government inducements.”); Mullock, *supra* note 11, at 750 (“[I]f we ask, ‘why did he do it?’ the answer is, ‘because he was predisposed to do it,’ and if we ask, ‘why was he predisposed to do it?’ the answer is, ‘because he did it.’”). While a defendant can, in theory, both deny committing the crime and

most obvious piece of evidence concerning the sort of person the defendant is, and some courts have found this sufficient evidence from which a jury can find predisposition.¹⁰⁴ Commission of the offense in the face of police inducement is, however, exceedingly minimal evidence of predisposition, given that predisposition is a measure of how the defendant would behave *in the absence of* police inducement. The circularity of this evidentiary bootstrapping is obvious when identified, but studies suggest that it may nonetheless tempt factfinders. The danger that the factfinder will be seduced by such a pernicious argument is magnified by the fundamental attribution error. In other words, to the same extent that the factfinder discounts the situational factors in the inducement inquiry, it is likely to overestimate the degree that the defendant's disposition played in his offense.¹⁰⁵

Richard Nisbett provided some of the clearest evidence of the propensity to excessively infer dispositional traits from manifestly situational behavior. Nisbett arranged to have some observer-subjects ("the observers") watch other actor-subjects ("the actors") participate in what they were told was a study on decisionmaking.¹⁰⁶ The observers watched as an experimenter announced to the actors awaiting the start of the experiment, "Before we get started, though, I happen to have sort of a real decision for you to make."¹⁰⁷ He told them that volunteers were needed to provide campus tours to potential donors to the University on the upcoming weekend.¹⁰⁸ If the actor was willing to volunteer, she would be paid for her time. One group of observers watched actors who were offered \$0.50 an hour,

argue that he was entrapped, *Mathews v. United States*, 485 U.S. 58 (1988), this strategy is rarely employed because the implication of inconsistency destroys the defendant's credibility. *See id.* at 67 (Scalia, J., concurring).

104. *See United States v. Tucker*, 28 F.3d 1420, 1429 (6th Cir. 1994) (upholding the district court's conclusion that a jury could find that the defendant was predisposed to commit the alleged crime because she committed the crime); *United States v. Jannotti*, 673 F.2d 578, 598-99 (3d Cir. 1982) (allowing jury to infer a politician's predisposition to accept bribes from his ready acceptance of the bribe in question); *Maestas v. United States*, 341 F.2d 493, 495 (10th Cir. 1965) (permitting a jury to find predisposition to sell narcotics from willingness to sell to undercover agent); *United States v. Myers*, 527 F. Supp. 1206, 1242 (E.D.N.Y. 1981) ("It is not unfair to permit a jury to infer a defendant's mental state, his predisposition, from the manner in which he responds to [the inducement].").

105. Consider, for example, the statement of Abscam informant Mel Weinberg that "a guy's either a crook or he isn't. If he ain't a crook, he ain't gonna do anything illegal no matter what I offer him or tell him to do." GARY T. MARX, *UNDERCOVER: POLICE SURVEILLANCE IN AMERICA* 120 (1988) (quoting Mel Weinberg).

106. Richard E. Nisbett et al., *Behavior as Seen by the Actor and as Seen by the Observer*, 27 J. PERSONALITY & SOC. PSYCHOL. 154 (1973); *see also* Jones & Nisbett, *supra* note 35 (discussing the study).

107. Nisbett et al., *supra* note 106, at 156.

108. *Id.*

while another group of observers watched actors who were offered \$1.50 an hour.¹⁰⁹ Only one-quarter of those offered \$0.50 volunteered, while two-thirds of those offered the higher wage accepted.¹¹⁰ As economic theory would predict, the wage, a situational factor, was a good predictor of whether the employment offer would be accepted.

Both the actors and the observers were later asked to predict the likelihood that the actor would volunteer to help another charitable group, the United Fund, for free.¹¹¹ Actors themselves did not think they were more likely to help the United Fund if they were volunteers than if they were nonvolunteers.¹¹² Observers, however, believed that if the actor had previously volunteered to give campus tours, she was much more likely to volunteer to help the United Fund than those who had not volunteered.¹¹³

Given that subjects were randomly assigned to the \$0.50 or \$1.50 group, there is little reason to think one group was more predisposed to volunteering. In making their predictions, however, observers of the high-volunteering, highly paid group judged them as generally being more likely to volunteer in the future than did the observers of the low-volunteering, poorly paid group.¹¹⁴ Observers apparently assumed that the actor's decision to volunteer reflected a predisposition to volunteer rather than the influence of the payment offered.¹¹⁵ When the actors were asked to explain why they accepted the offers, however, they were more likely to explain their behavior in terms of the amount of money they were offered,¹¹⁶ suggesting that people are much better at appreciating the situational causes of their own behavior than those of others.

The only factor distinguishing those who accepted the offer and those who did not was whether they were offered the higher or the lower amount of money — i.e. the degree to which they were induced to volunteer.¹¹⁷ But by predicting that actors would volunteer under less compelling circumstances, observers revealed that they largely

109. *Id.* Adjusted for inflation, these figures would be approximately \$2.07 and \$6.22 in 2003 dollars. See Consumer Price Index Inflation Calculator, at <http://www.jsc.nasa.gov/bu2/inflateCPI.html> (last visited Oct. 30, 2004).

110. Nisbett et al., *supra* note 106, at 157.

111. *Id.*

112. *Id.*

113. *Id.*

114. *Id.* at 157 tbl.1.

115. *Id.* at 157.

116. *Id.* (“It therefore appears that observers are inclined to make dispositional inferences from behavior under circumstances in which actors infer nothing about their general inclinations.”).

117. See *id.* at 157 tbl.1.

attributed the decision to volunteer to whether or not the actor had a volunteering disposition, and not the wage that he was offered.¹¹⁸ A jury confronted with a defendant pleading entrapment is similarly situated. The jury knows of an action — the crime — and a situation the actor was in — the inducement. If the conclusions of the Nisbett study hold true, the factfinder may infer a criminal predisposition from the induced crime.

Observers are also likely to infer an actor's disposition even when the actor is given no choice about his conduct. In a 1967 study, Edward Jones and Victor Harris asked subjects to read essays that either attacked or defended Castro's Cuba.¹¹⁹ One-half of the subjects were told that the author of the piece had been assigned a particular position to advocate for in a debate, with no choice on the author's part.¹²⁰ The other half were told that the author made his own choice of whether the essay would be pro- or anti-Castro.¹²¹ The subjects were then asked to estimate the author's actual opinion of Castro.¹²² Those who were told that the authors were given a choice largely believed, as expected, that those who wrote pro-Castro essays were truly pro-Castro and vice versa.¹²³

The assessments of those readers who were told that the author of their essay was given no choice, however, were quite surprising. Given that the subjects were told that the author was randomly assigned the editorial position for which he advocated, one would expect approximately similar attitudinal assessments of both pro- and anti-Castro authors. The editorial position of the essay, in other words, was irrelevant to the author's actual beliefs. Despite knowing that the authors had no choice concerning the position for which they advocated, this group of subjects judged the pro-Castro writers to be more pro-Castro than the anti-Castro writers.¹²⁴ The subjects gave substantial weight to the "face-value" meaning of the act of writing the

118. *Id.* at 157 ("It may be seen . . . that the actor's behavior prompted the observers to make dispositional inferences.")

119. Jones & Harris, *supra* note 34.

120. Jones and Harris explained the instructions to the subject concerning the essay as follows:

The essay was presented as the first draft of an opening statement in a college debate We assumed that the subjects would realize that debaters often try to defend positions in which they do not believe The choice-no choice manipulation was delivered orally: the debater had either been directed by the team advisor to argue a specified side of the topic or was given his choice of sides.

Id. at 8.

121. *Id.*

122. *Id.* at 4.

123. *Id.* at 6 tbl.1.

124. *Id.*

essay, severely failing to discount its implications in light of the obvious situational constraint.¹²⁵

The study suggests that even when a person has been assigned or coerced into a particular behavior, that behavior will be seen as indicative of his particular disposition.¹²⁶ The readers of the essay knew that the writers had no choice about the editorial position of their essays — the writers were perfectly constrained. Despite knowing this, the subjects still believed that the editorial position of the essay was indicative of the authors' beliefs.¹²⁷ Extrapolated into a courtroom setting, this suggests that even when the government inducement is so great as to leave no room for a person not to commit a crime, a jury is still likely to view the commission of the crime as probative of a criminal disposition.

2. *Failure to Apply Knowledge of Base Rates*

A postulate of classic attribution theory holds that causal explanations for an actor's behavior are or should be influenced by consensus information, i.e., information concerning how other people responded to a similar situation.¹²⁸ Situational factors are perceived as the primary cause of behavior when most people behave similarly, while the actor's unique disposition or preference is perceived as the cause when his behavior is unique.¹²⁹ Accordingly, we should expect that when a particular actor's behavior was the modal response to a particular inducement, that behavior is not highly probative of his disposition. A substantial amount of research has demonstrated,

125. *Id.* at 22 (“[Subjects] give substantial weight to the intrinsic or ‘face value’ meaning of the act itself in their attributions of attitude. This is true even when the act occurs in a no choice context.”).

126. Jones and Harris observed:

[C]orrespondence in attributing underlying attitudes to account for expressed opinions is high when the opinions are unexpected and expressed in a context of free choice. However, the content and direction of the opinions exert a clear inference on attribution even when choice is drastically reduced. In a context that permits the target person some very minimal degree of spontaneity, the perceiver seems to view his performance as more informative than a rational analysis of act and context would suggest.

Id. at 23.

127. *Id.* at 22. Follow-up studies have confirmed and elaborated on Jones and Harris's experiment. In one interesting formulation, experimenters asked subjects to rate how useful they found the essay in judging the writer's attitudes. Even those subjects who described the essay as “not useful” found the essay strongly probative of the writer's attitude. Arthur Miller et al., *The Perceived Value of Constrained Behavior: Pressures Toward Biased Inference in the Attitude Attribution Paradigm*, 47 *SOC. PSYCHOL. Q.* 160, 164 (1984). Other studies confirm that subjects make the same inferences when they read essays actually written by naïve subjects assigned to a particular position. Snyder & Jones, *supra* note 100.

128. *See, e.g.*, Kelley, *supra* note 33, at 208-09.

129. *Id.*

however, that knowledge of consensus information produces very weak effects on causal attributions.¹³⁰

In one study demonstrating this phenomenon, researchers described Milgram's obedience experiment to their subjects.¹³¹ One-half of the subjects were also told the results — that 65 percent of teachers administered the highest level of shock — while the other half were not.¹³² Both sets of subjects were then asked to rate various teachers who did or did not administer the maximum level of shock on a variety of character traits.¹³³ Subjects in both groups rated teachers who administered all of the shocks as weaker, colder, maladjusted, unattractive, unlikable, and more dependent than those who did not.¹³⁴

Data suggesting that obedience was largely induced had little bearing on the subjects' dispositional evaluations of the teachers.¹³⁵ Those subjects who knew that a majority of teachers completed the experiment did not substantially differ in the personality ratings they assigned.¹³⁶ Logically, being aware of this high base rate of obedience should decrease the correspondence between the act of shocking and a sadistic or aggressive disposition. In other words, if one knows that two-thirds of people behave in the same unexpected manner, one should be less prone to attribute dispositions as causes for such acts than if one assumes the behavior was aberrational.¹³⁷ But the results

130. See, e.g., Nisbett & Borgida, *supra* note 33; Amos Tversky & Daniel Kahneman, *Belief in the Law of Small Numbers*, 76 PSYCHOL. BULL. 105 (1971).

131. Arthur G. Miller et al., *Perception of Obedience to Authority*, in PROCEEDINGS OF THE 81ST ANNUAL CONVENTION OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION, 127, 127-28 (1972) [hereinafter Miller et al., *Perception*]; see also Arthur G. Miller et al., *The Prediction and Perception of Obedience to Authority*, 42 J. PERSONALITY 23 (1974) [hereinafter Miller et al., *Prediction*].

132. Miller et al., *Perception*, *supra* note 131, at 127.

133. *Id.*

134. *Id.* at 127-28. Milgram similarly noted:

Many people, not knowing much about the experiment, claim that subjects who go to the end of the board are sadistic. Nothing could be more foolish as an overall characterization of these persons. It is like saying that a person thrown into a swift-flowing stream is necessarily a fast swimmer, or that he has great stamina because he moves so rapidly relative to the bank. The context of action must always be considered.

Milgram, *Some Conditions*, *supra* note 75, at 72-73.

135. Miller et al., *Perception*, *supra* note 131, at 127; see also Martin A. Safer, *Attributing Evil to the Subject, Not the Situation: Student Reaction to Milgram's Film on Obedience*, 6 PERSONALITY & SOC. PSYCHOL. BULL. 205, 208 (1980) (suggesting that knowledge of the results of Milgram's experiment causes people to overestimate the aggressive characteristics of people generally).

136. Miller et al., *Perception*, *supra* note 131, at 127.

137. Professor Arthur Miller has noted:

[B]eing aware of this relatively high base rate of obedience should decrease the correspondence between the act of shocking the learner (particularly high shock) and personality and dispositions pertinent to such activity. Stated differently, if one knows that 65 percent of a sample of individuals perform in a somewhat unexpected manner, one should

suggest that knowledge that a particular behavior was the modal response to a situation will not substantially attenuate an observer's dispositional attributions.¹³⁸

Although a factfinder contemplating an entrapment defense would seldom be presented with evidence of what portion of the population would commit the charged offense given the police pressure, insensitivity to consensus information is relevant to the entrapment defense for two reasons. First, it confirms that the factfinder is not likely to make causal attributions or findings of predisposition in a manner that the scientist would describe as rational. Second, it suggests the robustness of observers' willingness to infer predisposition from situational behavior. The abstract, pallid nature of consensus information is unable to remedy the biased inferences produced by the more vivid, concrete commission of the offense.

C. *Overconfidence and Reasonable Doubt*

Determining inducement and predisposition is a difficult task, hampered by cognitive limitation. Were they aware of their own limitations, factfinders could mitigate the impact of these deficiencies by decreasing the confidence with which they make these findings, but research into the fundamental attribution error suggests that people are systematically overconfident in predicting others' behavior. Because the predisposition inquiry requires predicting whether the defendant was likely to commit the offense absent police pressure, factfinders are likely to be overconfident in making such findings. This overconfidence works peculiarly to the disadvantage of the defendant because it is he who is intended to enjoy the benefit of the reasonable doubt on the issue of entrapment.¹³⁹

be less prone to attribute personality dispositions as causes for such acts than if one assumes the behavior to be much less probable.

Miller et al., *Prediction*, *supra* note 131, at 39. Milgram similarly observed:

A commonly offered explanation is that those who shocked the victim at the most severe level were monsters, the sadistic fringe of society. But if one considers that almost two-thirds of the participants fall into the category of "obedient" subjects, and that they represented ordinary people . . . the argument becomes very shaky.

MILGRAM, OBEDIENCE TO AUTHORITY, *supra* note 75, at 5.

138. For further discussion of insensitivity to base rates in making predictions, see Daniel Kahneman & Amos Tversky, *On the Psychology of Prediction*, 80 *PSYCHOL. REV.* 237 (1973); Nisbett & Borgida, *supra* note 33; and Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 *SCIENCE* 1124 (1974).

139. See *Jacobson v. United States*, 503 U.S. 540, 548-49 (1992) ("Where the Government has induced an individual to break the law and the defense of entrapment is at issue . . . the prosecution must prove beyond reasonable doubt that the defendant was disposed to commit the criminal act prior to first being approached by Government agents.").

David Dunning, Lee Ross, Dale Griffin, and James Milojkovic demonstrated this phenomenon in a series of studies.¹⁴⁰ Each study required a subject to predict the response of a peer to various situations along with a “confidence estimate” that gauged the subject’s belief that his prediction would hold true.¹⁴¹ The amount of information the predictor had depended on which study he took part in, ranging from merely viewing a photograph of the actor to having been the actor’s roommate.¹⁴² The actor whose behavior was to be predicted was given a series of hypothetical situations and questions, and instructed to choose one of two potential responses.¹⁴³

The predictors consistently failed to achieve accuracy commensurate with their confidence levels.¹⁴⁴ In each version of the study, a clear majority of the subjects gave overconfident estimates. That is, average confidence estimates exceeded their average accuracy.¹⁴⁵ Subjects displayed overconfidence for 80% to 92% of the total predictions given.¹⁴⁶

The studies also demonstrated that increases in confidence outpace increases in accuracy.¹⁴⁷ The more confident a predictor was in an individual judgment, the more *overconfident* he was likely to be.¹⁴⁸ As a result, highly confident predictions tended to be highly overconfident predictions.¹⁴⁹

140. David Dunning et al., *The Overconfidence Effect in Social Prediction*, 58 J. PERSONALITY & SOC. PSYCHOL. 568 (1990).

141. *Id.* at 570.

142. *Id.* at 570-72.

143. *Id.* at 570. For example, if the actor found money on the floor, would he keep it or turn it in to the lost and found? *Id.* Other predictions included the target’s choice of magazine subscriptions between *Playboy* and the *New York Review of Books*, their customary mode of studying for an exam (studying alone v. study group), and the self-rated quality of their lecture notes (neat v. messy). *Id.*

144. *Id.* at 572.

145. *Id.*

146. *Id.* “Overconfidence” is defined as the subjects’ mean confidence estimate minus their mean accuracy. *Id.* For example, average confidence for predictions based on an interview or long-standing contact with a roommate ranged from 75% to 78% across the five studies, but average accuracy for those predictions was merely 60% to 68%. *Id.*

147. Low-confidence predictions were correct just over 50% of the time, medium-confidence predictions were correct about two-thirds of the time, and high-confidence predictions were correct about three-quarters of the time. *Id.* at 573. An accurate prediction was associated with a confidence rating of about 75%, and an inaccurate prediction with a confidence rating of 72%. *Id.*

148. *Id.* at 574.

149. *Id.* Predictions made with a low degree of confidence were on average overconfident by about six percentage points, and predictions made with a high degree of confidence were about fourteen to twenty-seven percentage points overconfident. *See id.* at 574 tbl.2. “High confidence” predictions were defined as a confidence rating of 90% or greater. *Id.*

Dunning and his colleagues offered two relevant explanations for overconfidence. First, virtually all of the overconfidence was due to against-base-rate predictions — that is subjects predicted the actor would behave differently from the modal response of his peers.¹⁵⁰ Indeed, the more pervasive the modal response, the more overconfident were predictions that the actor would deviate from it.¹⁵¹

The researchers' second explanation for overconfident behavioral prediction is that it results from overconfident situational construals.¹⁵² Predicting an actor or the average person's behavior in a given situation requires us to construe or even construct the incomplete situational data on which we rely. Dunning and his colleagues suggest, however, that predictors fail to adequately discount their confidence in their predictions when data concerning situational variables are uncertain. Instead, predictors typically generate a single construal when presented with limited data concerning an ambiguous situation, and then proceed to make predictions as if that construal corresponded to perfect knowledge.¹⁵³ The result is overconfident predictions and trait assessments.

In a follow up study, Dunning and his colleagues demonstrated the effect of construal processes on overconfident predictions.¹⁵⁴ Subjects were read a short description of a person's behavior in a given situation: Peter, a sophomore at Stanford, participated in a seventy-five minute, four-person discussion on abortion, during which he spoke for forty-five minutes.¹⁵⁵ The subjects were then asked to give an estimate of what portion of Stanford students would speak more than Peter in a similar group discussion, and provide a 50% confidence interval around this estimate.¹⁵⁶ They then assessed Peter according to three attributes: how opinionated, outgoing, and domineering he was relative to his peers; and they similarly provided confidence intervals for those traits.¹⁵⁷

The subjects were subdivided into four groups. The first was the "control group." Its members were merely given a chance to rethink

150. *Id.* at 576.

151. *Id.* at 577.

152. *Id.* at 579; see also Dale W. Griffin et al., *The Role of Construal Processes in Overconfident Predictions About the Self and Others*, 59 J. PERSONALITY & SOC. PSYCHOL. 1128 (1990).

153. Griffin et al., *supra* note 152, at 1138.

154. *Id.*

155. *Id.* at 1136.

156. *Id.* A 50% confidence interval corresponds to a range of times within which the subject believed the average Stanford student's speaking time would fall with 50% likelihood.

157. *Id.*

and alter their estimates as they saw fit.¹⁵⁸ The second was the “uncertain construal group.” Its members were asked to write a short essay describing how they construed the situation that Peter was in; and then, after they were finished, to provide a new set of trait assessment and confidence intervals.¹⁵⁹ The third was the “certain construal group,” and its members were similarly instructed to write an essay about how they construed Peter’s situation. Before remaking their assessments, however, the certain construal group was told to assume that the situation possessed the exact characteristics they just described in their essays.¹⁶⁰ Finally, the fourth group was the “multiple construal group.” Its members were instructed to write multiple essays about various ways the situation could have appeared, and were then asked to reassess their estimates.¹⁶¹ Judgments about both what portion of Stanford students would speak for a similar length of time and what character traits Peter possessed varied widely, reflecting the multiple reasonable construals of the ambiguous description of the group discussion.¹⁶²

Subjects in the control, uncertain, and certain construal groups all offered essentially the same sized confidence intervals before and after receiving their relevant construal instructions.¹⁶³ The researchers hypothesized that members of all three groups initially adopted a particular construal of Peter’s situation and made their estimates accordingly.¹⁶⁴

When the certain construal group was instructed to assume that their initial construal was in fact the correct one — to assume, in other words, that they now had perfect information about the situation — one would expect confidence to increase and the confidence interval to narrow. This was not the case. The addition of perfect information about the situational variables Peter was faced with had no significant effect on the certain construal group’s confidence.¹⁶⁵

158. *Id.*

159. *Id.*

160. *Id.* at 1131, 1136.

161. *Id.* at 1136.

162. *Id.*

163. *Id.*

164. *Id.* at 1138 (“These results support the more general contention that people typically generate a single construal of an ambiguous or incompletely specified situation and then, unless powerfully prompted to do otherwise, make relevant assessments and predictions as if their situational construals correspond to perfect knowledge.”).

165. *Id.* at 1136 (“Subjects in the certain construal condition, like subjects in the control and uncertain construal conditions, offered essentially the same size confidence intervals before and after receiving the relevant construal instructions.”).

The only group whose confidence interval changed from one estimation to the next was the multiple-construal group.¹⁶⁶ Members of that group were forced to brainstorm and consider reasonable situational construals other than their initial assessment.¹⁶⁷ When presented with the vagueness of their knowledge of the actual situation Peter faced, the multiple-construal group decreased their certainty in their estimations and widened their confidence intervals.¹⁶⁸ Only by making the situational ambiguity highly salient, by requiring subjects to generate alternate situational construals, was overconfidence reduced.¹⁶⁹

The researchers found that overconfidence in situational construals also produced extreme dispositional attributions.¹⁷⁰ Logically, uncertainty about the situation that produced Peter's behavior, especially given that his behavior seemed extreme, should produce a higher degree of conservatism in making trait inferences based solely on that single instance of behavior.¹⁷¹ When an actor's response seems extreme, we can either infer that the actor's response was indicative of his particular disposition, or we can assume that we

166. *Id.* at 1136.

167. *Id.*

168. *Id.*

169. *Id.* The researchers concluded:

We suggest that no matter how well one knows the particular individual — even if the individual to be predicted is the self — one will often be guilty of erroneous predictions if one fails to anticipate correctly what the details of the “situation” in question will actually be like and how the situation will be subjectively experienced. One will be guilty of overconfidence, furthermore, if one fails to recognize that such objective details and subjective representations matter a great deal . . . or fails to lower the subjective confidence of one's predictions in light of one's uncertainty about such details.

Id. at 1129. Professor Phoebe Ellsworth has similarly concluded:

Several different perceivers will come up with several somewhat dissimilar accounts of a sequence of events. Once having arrived at a construal, or a story or explanation of the same sequence of events, most people find it very difficult to imagine a different way of interpreting the same events, and this leads them to underestimate severely their own creative contribution to their “memory.” Even though most people recognize in principle that a good deal of perception is really interpretation, they are unable to make adequate inferential adjustments . . . often behaving exactly as they would if their interpretation were the only possible one.

Phoebe C. Ellsworth, *Are Twelve Heads Better than One?*, LAW & CONTEMP. PROBS., Autumn 1989, at 205, 206 (citing Griffin et al., *supra* note 152).

See also Robert P. Vallone et al., *Overconfident Prediction of Future Actions and Outcomes by Self and Others*, 58 J. PERSONALITY & SOC. PSYCHOL. 582 (1990) (discussing overconfidence in predictions of the subject's own behavior as compared to overconfidence in predicting the subject's roommate's behavior).

170. Griffin et al., *supra* note 152, at 1137.

171. *Id.* (“[U]ncertainty about the nature of the situation that prompted or provided the context for a given actor's response, especially when the response seemed extreme and potentially ‘diagnostic,’ should compel a rather high degree of conservatism in making trait inferences about such an actor.”).

erred in construing the situation and that the actor's behavior was not, after all, out of the ordinary. The more out of the ordinary the behavior seems, the more likely it was that our construal was inaccurate.

Dunning demonstrated, however, that subjects made little allowance for such situational uncertainty.¹⁷² Members of the certain construal group, for which there was no situational uncertainty, produced trait inferences no more extreme than the control or uncertain construal groups.¹⁷³ Members of the multiple-construal group, however, mitigated their dispositional inferences about Peter after considering other reasonable construals of the situation.¹⁷⁴

In an entrapment case applying the subjective standard, the prosecution bears the burden of proving beyond a reasonable doubt either that the defendant was predisposed or that he was not induced.¹⁷⁵ Given the difficulties judges and juries have in accurately making the factual findings that a plea of entrapment entails, we would expect them to have a high level of doubt and, accordingly, a high rate of acquittal. Experience indicates, however, that the entrapment defense is rarely successful.¹⁷⁶ Its inefficacy suggests that factfinders may not be accurately applying the reasonable doubt standard in entrapment cases. The low rate of acquittal may be explained by the effect of the fundamental attribution error, which can create overconfident causal attributions.

172. *Id.*

173. *Id.* ("As predicted, the certain construal condition subjects showed virtually no increase in the extremity of their trait inferences about Peter . . . and produced change scores that did not differ significantly from . . . the control and uncertain construal conditions . . .").

174. *Id.* Other studies confirm Dunning's conclusions. In one study by McGuire, observer-subjects were asked to predict the likelihood that actor-subjects would be helpful in two different situations. In the first, the actor was asked to volunteer for a psychology experiment, and in the second, the actor came upon a woman on crutches climbing some stairs whose book bag was about to slip off of her shoulder. For some of the observers, the subjects were unknown to them and only described to them in brief profiles; but in other conditions, the observers knew the actors quite well. The predictions were only slightly better than chance, but observers believed they would be quite accurate, especially if they knew the actor well. In reality, knowing the actor failed to significantly increase the accuracy of the observers' predictions. Observers were less accurate than they believed, and no more accurate when they had a great deal of knowledge about the person whose behavior they were trying to predict. The information they did have only served to make them overconfident in their prediction. ROSS & NISBETT, *supra* note 30, at 135 (describing A. McGuire, *Mistaken Reliance on Individual Difference Variables in Predicting Social Behavior* (1989) (unpublished manuscript, on file with the University of Michigan, Ann Arbor); see also Patricia G. Devine, *Overattribution Effect: The Role of Confidence and Attributional Complexity*, 52 SOC. PSYCHOL. Q. 149, 154 (1989) (providing confidence data for an experiment in the Jones and Harris framework).

175. *E.g.*, *United States v. Tom*, 330 F.3d 83, 89 (1st Cir. 2003).

176. See *supra* notes 20-25 and accompanying text.

This overconfidence works peculiarly to the disadvantage of the defendant in an entrapment case. When the factfinder determines whether the defendant is predisposed or not, it may consider a variety of character evidence and based on that, predict whether or not the defendant was likely to commit a similar offense on his own.¹⁷⁷ Assuming that most people are unlikely to commit a similar offense, a finding of predisposition would be “against the base rate”; it would be a finding that the defendant is likely to behave in a way different from most other people. The researchers found that against-the-base-rate predictions for a given actor are generally the most overconfident; and the more deviant the predicted behavior was, the more overconfident the prediction was.¹⁷⁸ Therefore, a factfinder who makes a determination that the defendant was predisposed is making an against-the-base-rate prediction and is likely to be overconfident in that factual finding. That overconfidence increases the likelihood that the judge or jury will find that predisposition was proven beyond a reasonable doubt, when a more accurate assessment of the evidence would reveal greater ambiguity. This effect is likely to be more pronounced the more unusual, and therefore more against-the-base-rate, the crime in question is.¹⁷⁹

It is tempting to conclude that increased uncertainty, whether justified or not, is neutral, benefiting neither party, but that would assume that the risk of error is shared equally between the prosecution and the defense. Overconfidence, however, works to the advantage of the party who has the burden of proof. In assessing the entrapment defense, the prosecution bears the burden of proving beyond a reasonable doubt that the defendant was predisposed.¹⁸⁰ Under this standard, the defendant should win whenever the factfinder believes that, more likely than not, the defendant was entrapped. The defendant should also win sometimes when the factfinder believes that the prosecution has proven that, more likely than not, the defendant was not entrapped, but has not met the more rigorous “beyond a reasonable doubt” standard. If the factfinder believes that, more likely than not, the defendant was entrapped, overconfidence has no effect. Regardless of whether the confidence in that decision is high or low,

177. *See supra* notes 7-11 and accompanying text.

178. *See supra* notes 170-174 and accompanying text.

179. This effect may be mitigated by the adversarial presentation of evidence. Insofar as this process compels the factfinder to consider both the prosecutor's and the defendant's interpretation of the sting operation, it requires consideration of at least two situational construals.

180. *Jacobson v. United States*, 503 U.S. 540, 549 (1992). Placing the burden of proof on the prosecution to rebut the defense represents a determination that “it is far worse to convict an innocent man than to let a guilty man go free.” *In re Winship*, 397 U.S. 358, 372 (1970) (Harlan, J., concurring).

the defendant is acquitted. But when the factfinder believes as an initial matter that the defendant was not entrapped, overconfidence works to the defendant's disadvantage by effectively lowering the burden of proof. If, for example, the factfinder should have a confidence level of only 60% that the defendant was not entrapped, the prosecution has not proven its case beyond a reasonable doubt.¹⁸¹ If Dunning's conclusions hold true, however, the perceived confidence level may exceed the reasonable-doubt threshold and result in an erroneous conviction.¹⁸²

In sum, studies in social prediction suggest that factual findings in an entrapment case are likely to be made with a higher degree of confidence than the evidence warrants. Insofar as these findings are affected by the fundamental attribution error, not only are they more likely to be incorrect, but they are more likely to be overconfident. Overconfidence in a finding that the defendant was predisposed based on the commission of the crime is likely worsened to the extent that factfinders fail to generate and contemplate multiple construals of the evidence concerning the police inducement. Finally, because the prosecution shoulders the burden of disproving the defense of entrapment beyond a reasonable doubt, overconfidence works to the defendant's disadvantage.

D. Channel Factors Applied to Sting Operations

Subjects have a hard time predicting the behavior of particular individuals or the average person in part because they underestimate the impact of certain situational aspects called "channel factors." A channel factor is a very small situational difference that produces surprisingly large changes in behavior.¹⁸³ In the study of the Good Samaritan, whether the seminarian was in a hurry produced a large change in behavior; therefore being rushed was a channel factor.¹⁸⁴

181. One could set the reasonable-doubt threshold at 91% based on Blackstone's famous statement that "the law holds, that it is better that ten guilty persons escape, than that one innocent suffer." 4 WILLIAM BLACKSTONE, COMMENTARIES 352 (photo. reprint 1992) (1765); cf. C.M.A. McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?*, 35 VAND. L. REV. 1293, 1332 (1982) (reporting that in a poll of 167 federal judges, the mean probability assigned to "beyond a reasonable doubt" was 90.28%). The reasoning of the argument contained in the text is unaffected by translating "beyond a reasonable doubt" into a less-exacting percentage, so long as it exceeds 50%.

182. I use the phrase "erroneous conviction" not in the sense that the defendant was convicted despite actually being entrapped. Rather, I use the term to mean that he was convicted despite a *substantial probability* that he was entrapped. That is, an erroneous conviction occurs when the prosecution did not prove the case beyond a reasonable doubt.

183. See generally ROSS & NISBETT, *supra* note 30, at 46-58 (discussing channel factors).

184. See *id.* at 48-49.

Experimenters have identified a variety of channel factors. In the Newton study of those most and least likely to donate to a food drive, receiving a map to the donation site, a personalized letter, and a phone call reminder were channel factors that greatly increased the likelihood that the recipient would donate.¹⁸⁵ Another commonly cited channel factor is the presence of a model — another person performing the action in question.¹⁸⁶ In subsequent experiments, Milgram showed that if teachers administered the learning experiment in a group setting and one teacher refused to continue, 90 percent of his fellow teachers followed suit.¹⁸⁷

One of the most potent channel factors, known as the “foot-in-the-door” technique, can be convincing a subject to take a small initial step along a path that would lead him to take much more substantial action. This phenomenon was aptly demonstrated by Jonathan Freedman and Scott Fraser in a classic study.¹⁸⁸ The researchers approached middle-class homemakers and requested that they take a small, innocuous step promoting a non-controversial cause, such as signing a petition or placing a small sticker in the corner of their car window supporting “safe driving.”¹⁸⁹ Two weeks later, a second researcher approached the homemakers, and also a control sample who had not previously been contacted, and requested that they take another more substantial step.¹⁹⁰ He asked them to place a large, poorly built, ugly “Drive Carefully” sign in their yard.¹⁹¹ Only 17 percent of the control group agreed to let the researchers place the sign, but 76 percent of those who had first signed the supporting petition did.¹⁹²

Channel factors have also proven useful in explaining previously puzzling behavior. For example, Lee Ross and Richard Nisbett have proposed the following channel factor based explanation of Milgram’s experiment on obedience.¹⁹³ First, the subjects formed an implicit contract with Milgram to complete the project and agreed to a specific

185. *See id.* at 132-33.

186. *See id.* at 49.

187. Milgram, *Some Conditions*, *supra* note 75, at 71.

188. Jonathan L. Freedman & Scott C. Fraser, *Compliance Without Pressure: The Foot-in-the-Door Technique*, 4 J. PERSONALITY & SOC. PSYCHOL. 195 (1966).

189. *Id.* at 199.

190. *Id.* at 200.

191. *Id.* (“The subject was shown a picture of a very large sign reading ‘Drive Carefully’ placed in front of an attractive house. The picture was taken so that the sign obscured much of the front of the house and completely concealed the doorway. It was rather poorly lettered.”).

192. *Id.* at 201 tbl.2.

193. ROSS & NISBETT, *supra* note 30, at 56-58.

procedure without realizing their full ramifications. The first few shocks were small, perhaps even innocuous. It was only through a series of gradual steps that subjects were led to administer dangerously high levels of voltage. If the experimenter had immediately instructed them to give the highest shock possible — one clearly labeled dangerous — he would likely have been widely disobeyed.¹⁹⁴ As one commentator has observed:

It is easy to see that there must be a line; it is not so easy to see where that line ought to be. . . . [I]f the subject decides that giving the next shock is not permissible . . . what was the justification for administering the last shock he just gave? . . . The subject is trapped by his gradual involvement in the experiment.¹⁹⁵

Until the learner begins thumping on the wall or stops responding, there is no clear point at which the subject can justify stopping now, as opposed to at the previous shock.¹⁹⁶ Consequently, it was at this point that most of the refusals to continue occurred.¹⁹⁷

Second, many of Milgram's subjects displayed a desire and intent to quit, but, in the absence of a channel factor through which that intent could be acted upon, most continued.¹⁹⁸ Ross and Nisbett suggest imagining that there was also a button on the control panel that the subject was told he could press whenever he wished to terminate the experiment.¹⁹⁹ They argue that this minor situational

194. Cf. Arnie Cann et al., *Effects of Initial Request Size and Timing of a Second Request on Compliance: The Foot in the Door and the Door in the Face*, 32 J. PERSONALITY & SOC. PSYCHOL. 774 (1975). Cann explored the effects of the onerousness of an initial request on the rate of acquiescence to a subsequent request. Cann asked one group of subjects to perform a relatively minimal task and then a task requiring an intermediate amount of time. He asked a second group of subjects to first perform a laborious task and then, after they responded, to perform the task requiring only an intermediate amount of time. Cann found that subjects in the second group were much less likely to comply. *Id.* at 777 tbl.2.

195. JOHN SABINI & MAURY SILVERS, *MORALITIES OF EVERYDAY LIFE* 70 (?); see John Sabini et al., *The Really Fundamental Attribution Error*, 12 PSYCHOL. INQUIRY 1, 3 (2001).

196. See also Sabini et al., *supra* note 195, at 3 (describing the “slippery slope” aspect of Milgram’s experiment as “crucial”).

197. ROSS & NISBETT, *supra* note 30, at 56.

198. Despite their high rate of compliance, many of Milgram’s subjects appeared quite eager to quit. Many were extremely distraught by their actions and displayed extreme tension, nervous laughter, trembling, and stuttering. Milgram, *Behavioral Study*, *supra* note 75, at 375. As one of Milgram’s associates stated:

I observed a mature and initially poised businessman enter the laboratory smiling and confident. Within 20 minutes he was reduced to a twitching, stuttering wreck, who was rapidly approaching a point of nervous collapse. He constantly pulled on his earlobe, and twisted his hands. At one point he pushed his fist into his forehead and muttered: “Oh God, let’s stop it.” And yet he continued to respond to every word of the experimenter, and obeyed to the end.

Id. at 377 (quoting an observer).

199. ROSS & NISBETT, *supra* note 30, at 57.

change would drastically increase the number of subjects who refuse to continue by providing a simple exit from an uncomfortable, inducing situation.²⁰⁰ One channel factor is, therefore, the provision of a means of escaping an inducing situation.

Third, Ross and Nisbett argue that the subjects were less likely to disobey the experimenter because the situation did not add up.²⁰¹ Although the subjects believed they were actually taking part in an experiment and actually administering severe shocks, the behavior of the experimenter was contrary to expectations. The experimenter did not express any concern about the safety of the learner or even check to make sure he was okay. In such a situation, where nothing makes sense, the subject may be less likely to act decisively or disavow role expectations.²⁰²

Milgram's follow-up study supports Ross and Nisbett's conclusion. Milgram repeated his experiment, but instead of requiring the teacher to incrementally increase the shock with each wrong answer, he allowed them to set any voltage level they wanted.²⁰³ Under those circumstances, only 2.5 percent of teachers administered the maximum shock.²⁰⁴ Milgram's original experiment was, therefore, not a particularly good measure of the teachers' disposition to inflict pain.²⁰⁵

A typical government sting is akin to a test for criminal predisposition.²⁰⁶ When the suspect takes the bait and commits the crime, he has tested positive. When the suspect declines the opportunity, he tests negative. A non-predisposed suspect who is induced into committing the crime is a "false-positive" and indicative of a failure in the sting's design. The desirability of a particular sting

200. *See id.*

201. *Id.* at 57-58.

202. Consider one teacher's response during post-experiment questioning:

My reactions were awfully peculiar. I don't know if you were watching me, but my reactions were giggly, and trying to stifle laughter. This isn't the way I usually am. This was a sheer reaction to a totally impossible situation. And my reaction was to the situation of having to hurt somebody. And being totally helpless and caught up in a set of circumstances where I just couldn't deviate and I couldn't try to help.

MILGRAM, OBEDIENCE TO AUTHORITY, *supra* note 75, at 54 (quoting a subject).

203. *Id.* at 70-72.

204. *Id.* at 60-61 tbl.3.

205. Milgram, however, never completely abandoned the search for dispositional explanations for why some teachers obeyed and why others disobeyed. *See id.* app. II. But, without using the term "channel factors," he described the teachers as "integrated into a situation that carries its own momentum. The subject's problem then is how to become disengaged from a situation which is moving in an altogether ugly direction." Milgram, *Some Conditions*, *supra* note 75, at 73.

206. *See Sorrells v. United States*, 287 U.S. 435, 441-42 (1932) ("The appropriate object of [a sting operation] frequently essential to the enforcement of the law, is to reveal the criminal design . . .").

design is largely a function of the number of “true-positives” — i.e., predisposed criminals — it catches and how few false-positives it produces. The entrapment defense, by acquitting some of the false-positives, is one way of discouraging the use of abusive sting operations and encouraging the police to more narrowly tailor their efforts.

Channel factors are one method that can be used to minimize false positives.²⁰⁷ For example, providing the defendant with a clear opportunity to terminate a high-pressure police encounter may protect the innocent. As discussed above, Ross and Nisbett hypothesized that many of Milgram’s teachers continued to shock learners for as long as they did because they did not perceive a clear means of exiting what was clearly an uncomfortable situation.²⁰⁸ The provision of an “experiment termination” button on the control panel likely would have substantially reduced the number of fully-compliant teachers.²⁰⁹ Similarly, Milgram demonstrated that fewer teachers administered the shocks when the experimenter issued commands by telephone instead of face-to-face.²¹⁰ Each of these measures would help, in Milgram’s experiment, to distinguish those truly predisposed to aggressive behavior from the ordinary person. Because they provide a method of separating the ordinary person from the predisposed, I shall refer to these situational manipulations as “exculpating channel factors.”

The goal of narrowly-tailored sting operations is, however, in tension with police incentives to maximize the number of convictions and the length of sentences.²¹¹ The police may, in other words, create

207. Another method of reducing false-positives is requiring the police to have reasonable suspicion that the target is engaged in or is preparing to engage in illegal activity before engaging in a sting operation. Teri L. Chambers, Note, *United States v. Jacobson: A Call for Reasonable Suspicion of Criminal Activity as a Threshold Limitation on Governmental Sting Operations*, 44 ARK. L. REV. 493, 510 (1991); J. Gregory Deis, Note, *Economics, Causation, and the Entrapment Defense*, 2001 U. ILL. L. REV. 1207, 1229 n.153; Jack B. Harrison, Note, *The Government as Pornographer: Government Sting Operations and Entrapment*: *United States v. Jacobson*, 916 F.2d 467 (8th Cir. 1990), rev’d, 112 S. Ct. 1535 (1992), 61 U. CIN. L. REV. 1067, 1088-94 (1993); Maura F.J. Whelan, Comment, *Lead Us Not into (Unwarranted) Temptation: A Proposal to Replace the Entrapment Defense with a Reasonable-Suspicion Requirement*, 133 U. PA. L. REV. 1193, 1216 (1985); Michael O. Zabriskie, Comment, *If the Postman Always “Stings” Twice, Who Is the Next Target? — An Examination of the Entrapment Theory*, 19 J. CONTEMP. L. 217, 237-44 (1993).

208. See *supra* Section I.D.

209. The signs of severe stress that most subjects displayed suggest that, if given a clear opportunity to escape, most subjects would have taken it. See *supra* notes 198-205 and accompanying text.

210. MILGRAM, OBEDIENCE TO AUTHORITY, *supra* note 75, at 59-62.

211. Cf. PATRICK M. WALL, EYE-WITNESS IDENTIFICATION IN CRIMINAL CASES 14 (1965). Wall argues:

identifications [of suspects] made by [eye-witness] policemen in highly competitive activities, such as undercover narcotic agents, whose chances for promotion may depend upon the number of arrests made because of their sales, should be scrutinized with special care. There is a danger that their identifications may be influenced unconsciously by their desire for

stings that maximize, rather than minimize, false-positives and encourage the already predisposed to commit more serious crimes than they otherwise would. One method of doing so is the use of what I will call “inculcating channel factors” — small situational changes that have the tendency to produce false-positives. The “foot-in-the-door” technique,” in which the police convince the defendant to commit some minor transgression in order to make him more agreeable to a larger scheme, is one such method especially prone to abuse.²¹² This could, for example, occur when a defendant is predisposed to sell small amounts of a drug, but not enough to qualify for a sufficiently substantial sentencing enhancement.²¹³ An informant or undercover officer could exploit this lesser disposition by establishing a small-scale drug operation with the suspect and then applying pressure to increase the quantity traded until the defendant qualified for the more severe sentence.²¹⁴

Use of inculcating channel factors poses a peculiar threat to the entrapment defense precisely because they can sharply effect behavior in ways that factfinders find difficult to predict and account for. Either through ignorance of the effects of channel factors, or willful manipulation, the police can design sting operations that produce a large number of false-positives that an unwary judge or jury will have

promotion.

Id. at 14.

212. *See supra* notes 188-192 and accompanying text.

213. Some courts have recognized a limited form of the defense known as “sentencing entrapment,” which occurs when the defendant is induced, through “outrageous official conduct,” to commit a more serious version of a crime than that to which he was predisposed. *See United States v. Si*, 343 F.3d 1116, 1128 (9th Cir. 2003); *United States v. Barth*, 990 F.2d 422, 424 (8th Cir. 1993). Sentencing entrapment may occur, for example, when a defendant, seeking to make a small drug buy, is pressured into purchasing a larger amount by an undercover agent. The burden of establishing sentencing entrapment is on the defendant. *United States v. Stavig*, 80 F.3d 1241, 1245 (8th Cir. 1996). A successful showing of sentencing entrapment does not result in acquittal, but merely a downward sentencing departure. *See, e.g., United States v. Staufer*, 38 F.3d 1103, 1108 (9th Cir. 1994) (finding sentencing entrapment and remanding for resentencing). The theory is not accepted by all the federal circuits, *see United States v. Miller*, 71 F.3d 813, 818 (11th Cir. 1996) (rejecting the theory of sentencing entrapment), and has been rejected by most state courts that have considered the issue. *See, e.g., Kelley v. State*, 821 So. 2d 1255, 1257 (Fla. Dist. Ct. App. 2002); *State v. Blackmon*, 78 S.W.3d 322, 332 (Tenn. Crim. App. 2001); *State v. Hardy*, 715 So. 2d 466, 472 (La. 1998); *Commonwealth v. Garcia*, 659 N.E.2d 741, 744 (Mass. 1996). *But see Leech v. State*, 66 P.3d 987, 989-90 (Okla. Crim. App. 2003) (adopting the theory); *Commonwealth v. Adams*, 760 A.2d 33, 40 (Pa. Super. Ct. 2000) (adopting same).

214. *See Leech v. State*, 66 P.3d 987 (Okla. Crim. App. 2003). In *Leech*, the defendant operated a small-scale methamphetamine operation with a government informant. *Id.* at 989. Their usual course of dealing was that the informant would borrow money from the defendant, which he would repay in small quantities of methamphetamine. *Id.* at 995 (Chapel, J., dissenting). On one occasion, the informant offered the defendant a quantity of methamphetamine worth 28 times the value of the debt as repayment. *Id.* at 995 n.31 (Chapel, J., dissenting). This amount also qualified the defendant for the more serious charge of trafficking. *Id.* at 995 (Chapel, J., dissenting).

difficulty separating from true-positives. Prosecutors, facing similar incentives to achieve convictions and lengthy sentences, and suffering from the same cognitive biases as factfinders, may not always be relied upon to refuse to prosecute a defendant snared in a heavy-handed sting.²¹⁵

The dangers that inculcating channel factors pose for the entrapment defense is illustrated by the First Circuit's decision in *United States v. Connell*.²¹⁶ In that case, an undercover agent arranged with a stockbroker, Connell, to launder money from a gambling operation in a series of transactions.²¹⁷ During the fourth transaction, the agent informed Connell that the money was actually derived from the illegal drug trade.²¹⁸ Knowledge or belief that the money being laundered was criminally derived carried with it a much stiffer sentence than laundering for other purposes, but the enhancement only applied when the defendant actually believed that the money was criminally derived, not when he "reasonably should have believed."²¹⁹ The purpose for the agent's deception was solely to expose Connell to a stiffer sentence. Accordingly, Connell argued that the sentencing enhancement should not apply because, although he was not entrapped into laundering money generally, he was entrapped into laundering money that he believed to be criminally derived.²²⁰

The police in *Connell* employed Freedman and Fraser's classic "foot-in-the-door" technique to ratchet up the severity of Connell's

215. One commentator observed after interviewing prosecutors:

In the extraordinary case where a possible [entrapment] defense is present . . . defense counsel will visit the prosecutor and argue that the charges should be dropped because the defendant was entrapped. The prosecutors never accede to this plea because of the lack of appeal the defense has to their sense of oughtness, because of their perception of their role in the criminal process, and because of their private ambitions.

It might be suspected that the resistance could be said to stem simply from the fact that "the law is on our side." But the prosecutors were found to have a distinct attitude toward the "entrapped defendant" in a narcotics sale case. They felt that his only excuse was faulty police methods, not a violation of the "letter" of the law, and the seriousness of his crime showed that he was not "innocent." While they all felt that the facts in *Toler* should have constituted entrapment, they all agreed that they would not have refused to prosecute the case unless it was shown "that the police officer twisted his [the defendant's] arm."

Bancroft, *supra* note 21, at 161 (footnote omitted). *People v. Toler* held that the defendant was not entrapped where the undercover agent requested drugs from him more than twenty times before he succumbed, and when the agent appealed to the defendant's sympathy by telling him the drugs were for a terminally ill narcotics addict. 185 N.E.2d 874 (Ill. 1962). Bancroft concluded that "the defense of entrapment *never* triggers the decision not to prosecute." Bancroft, *supra* note 21, at 162 (emphasis added).

216. 960 F.2d 191 (1st Cir. 1992).

217. *Id.* at 193.

218. *Id.*

219. *Id.* at 195 n.7.

220. *Id.* at 194.

sentence.²²¹ The police and Connell agreed to a crime that Connell was predisposed to — laundering non-drug derived money. By exposing himself to this degree of criminal liability, he became personally invested in the criminal enterprise. The subsequent request to launder drug money was — other than the money's origin — a request that differed only in degree but not in kind.

A comparison to Milgram's teachers is illustrative. Both Connell and the teachers agreed to an initial set of procedures.²²² In Connell's case, it was money laundering for a non-criminal enterprise.²²³ Although this act was illegal, it does not compel a conclusion that he was predisposed to launder money for drug dealers, but neither Milgram's teachers nor Connell realized the scope of the procedures to which they had agreed. Both Milgram's teachers and Connell engaged in a series of steps in furtherance of the arrangement and then were presented with a critical juncture. For the teachers, it was the learner banging on the wall and ceasing to respond to questions; for Connell, it was being told that he had become part of a major drug operation.²²⁴ Presumably, the police believed that Connell was less likely to agree to launder money for a drug operation; that is why they waited until he had conducted three illegal transactions before telling him. Similarly, Milgram knew that teacher obedience would be substantially reduced if he asked them to start off giving the highest level of shock to a protesting learner. And neither the teachers nor Connell were provided with an exculpating channel factor. It seems unlikely that a stock broker, when confronted with the fact that he is laundering money for a ring of drug dealers, would feel free to call the arrangement off in the absence of a clear escape route. Similarly, Milgram's subjects were incapable of putting their generalized desire to quit administering shocks into action in the absence of a clear means of acting on it.

The court rejected Connell's sentencing entrapment argument, reasoning:

221. See *supra* Section I.D.

222. Compare *Connell*, 960 F.2d at 193 (describing the procedures that Connell and the police agreed upon for their money laundering operation), with *Milgram, Behavioral Study*, *supra* note 75, at 373 (describing the procedure for Milgram's experiment to which the subjects agreed).

223. *Connell*, 960 F.2d at 193 (“During their first meeting, [the undercover agent] told Connell that the money was coming from an elaborate gambling operation in Atlantic City (whether legal or illegal, [the agent] did not specify).”).

224. *Id.* at 194.

[Connell] contends that the vice lay in the timing: by broaching the subject of the currency's supposed origin (drug trafficking) only after Connell had fully completed three episodes of money laundering, the undercover agent forced (or lured) him into actions he would otherwise have eschewed, i.e., peripheral participation in the narcotics trade.

Id.

By their nature, sting operations are designed to tempt the criminally inclined, and a well-constructed sting is often sculpted to test the limits of the target's criminal inclinations. Courts should go very slowly before staking out rules that will deter government agents from the proper performance of their investigative duties.²²⁵

To the contrary, because a sting operation is intended to merely produce a temporal displacement of crime — causing its commission at a time when the police can catch the defendant — a well-designed sting operation should not strive to push the defendant beyond the limits he would have faced in his natural environment. “[T]est[ing] the limits of the target's criminal inclinations”²²⁶ when he would not have been so tested otherwise is to engage in the inefficient, sterile activity of first inciting crime and then punishing, which the entrapment defense was meant to prevent.²²⁷ As Milgram's experiment proves, most people have some potential to commit frightening, criminal acts.²²⁸ In response to adept manipulation of channel factors, the ordinary person may commit criminal acts well beyond what would ordinarily be expected. If the channel factor is subtle enough, the police conduct can dramatically affect behavior while never approaching what a judge is likely to perceive as sufficiently egregious to warrant a sentencing entrapment defense.

It is impossible to know how Connell would have behaved under a more narrowly tailored sting operation. Whether he would have declined to launder drug money if it had been the first request that the undercover agent had made, or whether he would have quit had he been told he could, is speculation. Whichever the case, the *Connell* court's statement that “we find no grounds for concern in the circumstances at bar”²²⁹ displays a troubling lack of consideration of both how the police can use situational variables to sculpt a defendant's behavior, and how a tailored sting operation could be conducted to minimize false positives.²³⁰ Closer attention in future cases to whether the government took efforts to offer — or steps to

225. *Id.* at 196.

226. *Id.*

227. *See* United States v. Kaminski, 703 F.2d 1004, 1010 (7th Cir. 1983) (Posner, J., concurring) (arguing that one purpose of the entrapment defense is to prevent wasting law enforcement resources by instigating crime that would not otherwise occur).

228. Had the learner in fact been shocked, many versions of Milgram's experiment would have required the teacher to commit criminal acts. *See, e.g.,* MILGRAM, OBEDIENCE TO AUTHORITY, *supra* note 75, at 3-4 (describing experiment and fact that, under some variations, the learner demanded to be released and the teacher was required to forcibly place the learner's hand on a shock plate).

229. *Connell*, 960 F.2d at 196.

230. The FBI, for example, has promulgated guidelines that instruct agents to model undercover operations on the real world as closely as they can. MARX, *supra* note 105, at 182.

avoid offering — exculpating channel factors could focus judicial scrutiny on the situational nuance that police officers potentially exploit.

II. SOLUTIONS AND ALTERNATIVES

This Note has argued that the fundamental attribution error produces factfinding inaccuracies that diminish the ability of the entrapment defense to exculpate the unwary innocent and serve as a meaningful check on police overreaching. In addition, systematic biases in factfinding are ripe for exploitation by overzealous law enforcement who could, for example, tailor sting operations to maximize the number of unprejudiced targets who commit the offense, thereby maximizing arrests, convictions, and sentences at the expense of accuracy.

This Part proposes three modest reforms that might remedy some of the dangers the fundamental attribution error poses for the entrapment defense. First, it argues that the objective test should be adopted as a question of fact for the jury. Second, it proposes that, at a minimum, some additional instruction to the jury can attenuate some of the effects of the fundamental attribution error. Finally, it suggests that the failure of the police to provide exculpating channel factors and the inclusion of inculcating channel factors in the design of sting operations should be looked upon with greater suspicion by the courts.

A. *An Objective Test Decided by a Jury*

The subjective test for entrapment, as applied in the federal courts, classifies both predisposition and inducement as questions of fact for the jury. The objective test in most jurisdictions differs in two regards. First, the objective test is applied by a judge, not a jury.²³¹ Second, it focuses solely on the degree of inducement offered by the police, not the defendant's predisposition.²³² Common reasons for preferring that the judge decide the issue are that he is better qualified to set standards for future police conduct,²³³ jury verdicts give little future guidance for sting operations,²³⁴ the judge is less likely to be swayed by

231. The question-of-law approach has been articulated in the Supreme Court by Justice Roberts, concurring in *Sorrels v. United States*, 287 U.S. 435, 457 (1932) (Roberts, J., concurring) ("It is the province of the court and the court alone to protect itself and the government from such prostitution of the criminal law. . . . Proof of entrapment, at any stage of the case, requires the court to stop the prosecution, direct that the indictment be quashed, and the defendant set at liberty.")

232. See, e.g., UTAH CODE ANN. § 76-2-303 (2003); MODEL PENAL CODE § 2.13 (1962).

233. MARCUS, *supra* note 3, at 189.

234. Judge Traynor, for example, has stated:

otherwise-inadmissible evidence of prior bad acts, or that “evidence pertaining to guilt is likely to infect a jury determination.”²³⁵

This Section takes a different, somewhat unconventional approach,²³⁶ adopting some of the arguments others have made in favor of the objective test, but contending that the jury should apply it, not a judge. First, entrapment should be a question for the jury, which is more likely to engage in group discussion, an activity that has been shown to attenuate the effects of the fundamental attribution error. Second, the jury should apply the objective standard because it eliminates the focus on the defendant’s predisposition, the inquiry most influenced by a dispositional bias and the use of otherwise-inadmissible evidence of prior bad acts.²³⁷

1. *The Benefits of Jury Discussion*

The proper division of labor between the judge and jury has been the source of extensive contemporary debate.²³⁸ I do not seek to

A jury verdict of guilty or not guilty tells the police nothing about the jury’s evaluation of the police conduct. A verdict of guilty may mean that the jury did not believe the defendant’s testimony that would have established entrapment. It may also mean that the jury did not believe that the conduct created a substantial risk of inducing one not ready to commit the offense into doing so.

People v. Moran, 463 P.2d 763, 769 (Cal. 1970) (Traynor, C.J., dissenting).

235. People v. D’Angelo, 257 N.W.2d 655, 659 (Mich. 1977).

236. See, e.g., WAYNE R. LAFAYE ET AL., CRIMINAL PROCEDURE § 5.3(b) (stating that “it is not entirely clear why [the objective test should be a jury question]”); Myron Moskowitz, *You Can’t Tell a Book by Its Title*, 8 CRIM. L.F. 125, 134-35 (1997) (“I can’t see why entrapment should ever be decided by a jury.”) (reviewing H. RICHARD UVILLER, VIRTUAL JUSTICE: THE FLAWED PROSECUTION OF CRIME IN AMERICA (1996)); Laura Gardner Webster, *Building a Better Mousetrap: Reconstructing Federal Entrapment Theory from Sorrells to Mathews*, 32 ARIZ. L. REV. 605, 630 (1990) (describing the “breathtaking naivete” with which the Supreme Court allows the jury to decide entrapment).

237. It should be cautioned, however, that although a jury-determined objective test may improve the accuracy of factfinding, juries may have more difficulty understanding the law of the objective test than the subjective test. See Eugene Borgida & Roger Park, *The Entrapment Defense: Juror Comprehension and Decision Making*, 12 LAW & HUM. BEHAV. 19 (1988) (arguing that juror comprehension of the objective standard was lower than that of the subjective standard). This, however, may be the result not of any inherent conceptual complexity within the objective standard, but with overly confusing ways of explaining it in jury instructions. See *id.* at 35-36 & n.15. Subjects in psychological studies, for example, do not seem to have difficulty understanding what it means to predict what an “average person” will do in a given situation. But it is unsurprising that a juror might not understand what it means to determine what a “hypothetical, law-abiding, non-predisposed, reasonable person” would do in response to police pressure. The problem with juror comprehension of the objective test may not lie, therefore, in the form of the test, but in the form of the instruction.

238. On the question of whether the judge or jury is a better factfinder, see Ellsworth, *supra* note 169, at 217-18; Richard Lempert, *Civil Juries and Complex Cases: Taking Stock after Twelve Years*, in VERDICT: ASSESSING THE CIVIL JURY SYSTEM 181 (Robert Litan ed., 1993); Robert MacCoun, *Inside the Black Box: What Empirical Research Tells Us about Decisionmaking by Civil Juries*, in VERDICT: ASSESSING THE CIVIL JURY SYSTEM 137

resolve the dispute, but instead wish to suggest one reason for preferring that juries resolve the issue of entrapment: juries enjoy the benefit of group discussion.

A series of studies suggests that the process of deliberation and discussion ameliorates the effects of the fundamental attribution error. In one study, Edward Wright and Gary Wells replicated a scenario very similar to Jones and Harris's study of the readers of pro- and anti-Castro essays.²³⁹ That is, all subjects read essays arguing a particular position. One half of the readers were told the writer was assigned his editorial position, while the other half were told that he chose it.²⁴⁰ Wright and Wells also divided the readers of the essays into four groups: those who would immediately answer questions about the essay's author after reading it, those who would answer the questions after a ten minute delay, those who were given a ten minute delay and told they would discuss their answers with a group after answering, and those who were actually given ten minutes to discuss the questions with a group before answering.²⁴¹

The results indicated that group discussion substantially reduced the impact of the fundamental attribution error.²⁴² Those subjects who engaged in group discussion were much less likely to erroneously discern the author's dispositional traits from the editorial position of an essay that he had no choice in deciding.²⁴³ Neither anticipating that one would engage in group discussion after answering the questionnaire nor having an additional ten minutes to think about the questionnaire before answering had a significant effect.²⁴⁴ Those who participated in group discussion were much less likely to draw dispositional inferences from situationally constrained behavior, but Wright and Wells were not able to provide a definitive explanation of

(Robert Litan ed., 1993); *see also* United States v. Booker, 125 S.Ct. 738 (2005) (holding that the mandatory nature of the Federal Sentencing Guidelines violated the Sixth Amendment due to judicial factfinding); Blakely v. Washington, 124 S. Ct. 2531 (2004) (holding that the maximum sentence a judge may impose is based on the facts admitted by the defendant or found by a jury, not the judge).

239. Edward F. Wright & Gary L. Wells, *Does Group Discussion Attenuate the Dispositional Bias?*, 15 J. APPLIED SOC. PSYCHOL. 531 (1985). Wells, however, instructed the authors to write about whether a portion of Canada should secede. *Id.* at 535.

240. *Id.* at 536-37.

241. *Id.* at 536.

242. *Id.* at 542 ("The tendency of perceivers to make dispositionally-biased attributions for behavior performed under constraint was not only reduced, but eliminated, when attributors were instructed to render their judgment after a group discussion of the critical attribution question.").

243. *Id.* at 540 tbl.1. Recall that because the readers were told that the editorial position of the essay was assigned to the writer and not of his own volition, it was logically irrelevant to the writer's actual opinion.

244. *Id.*

the phenomenon.²⁴⁵ Subsequent studies have confirmed the value of group discussion in making dispositional attributions.²⁴⁶

A plausible explanation for the benefits of group discussion may be found in the overconfidence studies of David Dunning.²⁴⁷ Recall that when Dunning forced subjects to write essays providing multiple interpretations of an ambiguous set of facts, the effects of the fundamental attribution error were reduced. Subjects were less likely to infer dispositional traits from situationally constrained behavior, and they were less likely to be overconfident. The act of writing essays in Dunning's experiment may have had the same effect as the group discussion in the Wells and Wright experiment. Just as writing essays providing alternate construals of an ambiguous situation forced Dunning's subjects to confront the uncertainty of their initial assessment of the facts, so might discussing those facts with a diverse group of strangers with unique perspectives. Although this may not result in a complete revision of their interpretation, it reminds jurors that their construal of the facts is not always the only reasonable one.²⁴⁸

The work of Wells, Wright, and Dunning therefore provides an overlooked reason to allow the jury to decide the question of entrapment, as it commonly does under the subjective test. The objective test as it is commonly applied, is flawed in that it views the entrapment inquiry as exclusively a question of law for the judge.²⁴⁹ There is, however, no guarantee that a judge would engage in fundamental-attribution-error-attenuating group discussion. Indeed, unless one of her law clerks observed all of the relevant testimony, there is no one with whom she could effectively do so. Those jurisdictions that reserve the entrapment inquiry for the judge abandon the best-known method of reducing the effects of the fundamental attribution error: discussion.²⁵⁰

245. *Id.* at 544.

246. See Gwen M. Wittenbaum & Garold Stasser, *The Role of Prior Expectancy and Group Discussion in the Attribution of Attitudes*, 31 J. EXPERIMENTAL SOC. PSYCHOL. 82 (1995). As encouraging as these studies are, the benefits of group discussion do not eliminate the effects of the fundamental attribution error, especially when the inquiry is one of judgment, without a demonstrably correct answer. *Id.* at 102.

247. See *supra* Section I.C.

248. See Ellsworth, *supra* note 169, at 206 ("If it does nothing else, group deliberation . . . forces people to realize that there are different ways of interpreting the same facts A judge does not have this vivid reminder that alternative construals are possible."). The adversary process by itself may by itself encourage the consideration of multiple construals, but more research is needed to determine to what extent.

249. See, e.g., *State v. Valdez-Molina*, 897 P.2d 993, 995 (Idaho 1995); *People v. D'Angelo*, 257 N.W.2d 655, 660 (Mich. 1977).

250. This is not to suggest either that the average judge is worse than the average lone juror, or that the exceptional judge might not be less prone to the fundamental attribution error than the average jury. Indeed, studies have shown that certain individuals who score

2. *The Benefits of the Objective Test*

The objective test is preferable on two grounds related to the accuracy of factfinding. First, because it focuses solely on the inducing effects of the police conduct, while abandoning any inquiry into the defendant's predisposition, it eliminates the element of the subjective test that is most susceptible to bias by the fundamental attribution error. Second, because the disposition of the defendant is no longer relevant, there is no need to admit prejudicial character and bad-acts evidence.²⁵¹

Studies suggest that the effects that the fundamental attribution error has on the predisposition inquiry are more pervasive and more difficult to remedy. In one illuminating experiment, Günter Bierbrauer demonstrated that subjects placed under certain conditions are better able to predict the behavior of the ordinary "teacher" in Milgram's experiment.²⁵² Bierbrauer recreated Milgram's experiment, and allowed subjects to observe one teacher administer the entire sequence of shocks.²⁵³ After witnessing the experiment, subjects were asked to assess both the inducing effect of the experiment and to make dispositional attributions to the particular teacher they observed.²⁵⁴ Specifically, they were required to predict the percentage of teachers who would refuse to continue at various shock levels, and the highest level of shock that the subject himself, his best friend, the learner, and the average Stanford student would administer if they were the teacher.²⁵⁵ They were then asked to rate the teacher they had just observed on four personality traits and predict his behavior in five hypothetical situations.²⁵⁶ One-third of Bierbrauer's subjects were required to answer these questions immediately after witnessing the

well on the "Attributional Complexity Scale," see Garth J.O. Fletcher et al., *Attributional Complexity: An Individual Differences Measure*, 51 J. PERSONALITY & SOC. PSYCHOL. 875 (1986), are less prone to the fundamental attribution error. See Devine, *supra* note 174.

One possible compromise solution would be to give the defendant the option of trying the question of entrapment to a jury at trial, or to a judge before trial. Cf. *State v. Grilli*, 230 N.W.2d 445, 455 (Minn. 1975) (applying a subjective entrapment theory and granting the defendant the option to present the defense to either a judge in a pre-trial hearing, or a jury at trial).

251. See FED. R. EVID. 404 (prohibiting the use of character evidence and evidence of other crimes "to show action in conformity therewith," but allowing admission for other purposes, such as proving predisposition).

252. Günter Bierbrauer, *Why Did He Do It? Attribution of Obedience and the Phenomenon of Dispositional Bias*, 9 EUR. J. SOC. PSYCHOL. 67, 73 tbl.1 (1979).

253. *Id.* at 71.

254. *Id.* at 72.

255. *Id.* at 72.

256. *Id.* at 72-73. The hypothetical situations involved the teacher "making decisions involving adaptability, life-rescuing, compliance, cheating, and personal interference." *Id.* at 73.

experiment, one-third were required to play a number of games for thirty minutes before answering, and one-third were instructed to write and think about the experiment for thirty minutes before answering.²⁵⁷

The results revealed that delaying the subjects' responses improved their perception of situational control. Those subjects who were instructed to write and think about the experiment performed the best, predicting that the average student would administer a higher maximum shock and that a smaller percentage would disobey.²⁵⁸ Although delayed-response subjects still predicted higher rates of disobedience than Milgram demonstrated, time for contemplation drastically improved the subjects' assessment of how the average person would perform in Milgram's experiment.²⁵⁹ Thinking about situational constraints, in other words, increased the accuracy of the subjects' assessment of the inducing effects of Milgram's experiment.

Bierbrauer's subjects, however, did not seem to benefit from contemplation when it came to inquiries into what more closely resembled predisposition. Subjects in all three groups drew strong dispositional inferences about the teacher they witnessed, and they were willing to make strong predictions about his future behavior in hypothetical situations.²⁶⁰ There was no statistically significant difference in this regard between the three groups.²⁶¹ Bierbrauer concluded that "opportunity to contemplate the witnessed behaviour does not decrease dispositional attribution."²⁶² This suggests that the process of deliberation may increase the accuracy of the inducement inquiry (what the ordinary person will do) somewhat, but fails to increase the accuracy of the predisposition inquiry (what qualities the defendant has or whether he is likely to break the law absent police persuasion).

The effects of the fundamental attribution error might be further attenuated by instruction that encourages additional deliberation. Merely anticipating future discussion does not seem to provide the same benefits as the actual process of group discussion.²⁶³ Approximately one-half of all juries begin their deliberations with a vote.²⁶⁴ In addition to the danger that an opening vote will commit

257. *Id.* at 71.

258. *Id.* at 73 tbl.1.

259. *Id.*

260. *Id.*

261. *Id.* at 75-76.

262. *Id.* at 76.

263. Wright & Wells, *supra* note 239, at 542.

264. Ellsworth, *supra* note 169, at 214.

jurors to their initial position too early, the jury might also achieve discussion-obviating unanimity. Because the amount of dialogue required to substantially diminish the impact of the fundamental attribution error is quite small, measures that encourage at least a minimal period of discussion, such as discouraging the jury from beginning deliberation with a vote, are a virtually costless method of increasing jury accuracy.²⁶⁵

The second and more obvious benefit from eliminating the predisposition element is that the defendant's predisposition is no longer relevant, thereby eliminating the need to introduce prejudicial character evidence.²⁶⁶ When courts apply the subjective test for entrapment, they routinely admit evidence relevant to the defendant's predisposition that would, under normal circumstances, be excluded as impermissible character evidence, such as evidence of the defendant's prior bad acts.²⁶⁷ Introduction of such evidence invites the jury to convict the defendant based upon her prior, uncharged conduct.²⁶⁸ Adoption of the objective test eliminates this danger.

B. *Closing Arguments*

The easiest, least expensive, and least objectionable partial remedy for these problems is a more fully informed jury. The factfinding accuracy of the jury may be improved through a variety of techniques. It may, for example, help if jurors are encouraged to generate multiple situational construals of the evidence.

Closing arguments may also suggest that the jurors imagine themselves in the defendant's situation. Studies suggest that experiencing the situational constraints a subject was under when she acted tends to mitigate the effects of the fundamental attribution error. In their experiment on attitude attribution from pro- and anti-Castro essays, Jones and Harris discovered that one way they could reduce their subjects' likelihood of drawing unwarranted inferences was by requiring them to write their own assigned-position pro- or

265. See Wright & Wells, *supra* note 239, at 542 (concluding that increased decision *time* did not further the disposition-attenuating effects of discussion so long as some discussion occurred).

266. *Sherman v. United States*, 356 U.S. 369, 383 (1958) (Frankfurter, J., concurring).

267. See, e.g., *Sorrells v. United States*, 287 U.S. 435, 451 (1932) (stating that "if the defendant seeks acquittal by reason of entrapment he cannot complain of an appropriate and searching inquiry into his own conduct and predisposition as bearing upon that issue").

268. D. Craig Lewis, *Proof and Prejudice: A Constitutional Challenge to the Treatment of Prejudicial Evidence in Federal Criminal Cases*, 64 WASH. L. REV. 289, 325 (1989) ("[E]vidence of uncharged misconduct can lead a jury to convict an accused, even if guilt of the charged offense has not been clearly demonstrated, because the accused has been shown either to be a person deserving of punishment for bad character or to be guilty of other sins for which the accused has never been punished.").

anti-Castro essay before judging the disposition of other authors.²⁶⁹ When the salience of the situational constraint was increased, the subjects were more hesitant to assume the disposition of others similarly situated.²⁷⁰ This, others have argued, is because actors are much more likely to see their actions as situationally constrained.²⁷¹ The actor's attention is focused outward on situational cues, rather than inward on her own behavior, but for the observer, the focal stimulus is the actor himself.²⁷²

The nature of the adversarial process already provides a substantial opportunity to present alternative interpretations of the facts surrounding the commission of an offense. This effect can be amplified by jury instruction or defense closing argument that encourages the jury to imagine themselves in the role of the defendant and how the juror would have explained the causes of her behavior in that situation. If the jurors are asked to imagine themselves in the defendant's position, it may help them to appreciate the situational pressures she was under.²⁷³

269. See Jones & Harris, *supra* note 34, at 10 tbl.2; see also Sheldon Ungar & Aysan Se'ver, "Say It Ain't So, Ben": *Attributions for a Fallen Hero*, 52 SOC. PSYCHOL. Q. 207 (1989) (hypothesizing and finding evidence for the proposition that when subjects identify with an actor, and the actor does something disreputable, they are more likely to attribute his behavior to situational factors).

270. See Jones & Harris, *supra* note 34, at 12 ("When the prediction task was preceded by the task of writing a pro-Castro speech under directions, the correlation vanishes. Having to write a speech against one's own position seems to reduce the significance of that position when it comes to imputing the attitude of a target person operating under the same prescription."). This effect has been replicated when subjects are forced to observe their own behavior from a third-person perspective, for example through a videotape of their actions. See Michael D. Storms, *Videotape and the Attribution Process: Reversing Actors' and Observers' Points of View*, 27 J. PERSONALITY & SOC. PSYCHOL. 165 (1973).

271. Jones & Nisbett, *supra* note 35.

272. *Id.* at 7. For a recent recreation and expansion of the Jones & Harris study, see Shiri Nussbaum et al., *Creeping Dispositionism: The Temporal Dynamics of Behavior Prediction*, 84 J. PERSONALITY & SOC. PSYCHOL. 485 (2003).

273. It may be argued that this would be a prohibited "golden rule" argument, which asks jurors to place themselves in the position of one of the parties. See, e.g., *State v. McHenry*, 78 P.3d 403, 410 (Kan. 2003); *Forbes v. State*, 771 So.2d 942, 950 (Miss. Ct. App. 2000); *Gomez v. State*, 751 So.2d 630, 632 (Fla. Dist. Ct. App. 1999); *King v. State*, 877 S.W.2d 583, 586 (Ark. 1994); *Chisolm v. State*, 529 So.2d 635, 640 (Miss. 1988). The rationale for the prohibition is that such an argument, by in effect asking jurors to "do unto" one of the parties what they would have that party "do unto them," encourages the jury to depart from neutrality. *Loose v. Offshore Navigation, Inc.*, 670 F.2d 493, 496 (5th Cir. 1982). The prohibition on golden rule arguments does not, however, apply when used to ask the jury to assess the reasonableness of a party's actions. See *Cummins Ala., Inc., v. Allbritten*, 548 So.2d 258, 263 (Fla. Dist. App. 1989) (allowing a golden rule argument in a negligence case which asked what precautions the jurors, as reasonable people, would have taken). Golden-rule-type jury instructions are also given when a defendant pleads self-defense, which requires the jury to assess the reasonableness of the defendant's fear of bodily harm. See, e.g., *Hood v. State*, 27 So. 643, 644 (Miss. 1900) (finding reversible error where trial court refused a proposed self-defense jury instruction containing the statement, "[t]he jury must put themselves, as far as possible, in the defendant Hood's place, and then judge whether the danger was apparent"); 4 OHIO JURY INSTRUCTIONS CRIM. 411.35 § 2 (2002) ("In deciding

C. *Consideration of Inculcating and Exculpating Channel Factors in Motions for a Finding of Entrapment as a Matter of Law*

The consideration of channel factors provides one valuable technique for deciding whether entrapment has been established as a matter of law. In particular, judges can inform themselves of commonly used inculcating and exculpating channel factors, and use that knowledge when deciding whether to allow the question of inducement to go to the jury. Should a judge identify the use of inculcating channel factors, the question of inducement should be given to the jury only with hesitancy. If, however, the police provided an exculpating channel factor, a judge can be confident that inducement is appropriately a question for the jury.

The Supreme Court hinted at such an approach in *Masciale v. United States*.²⁷⁴ In *Masciale*, the defendant was convicted of selling heroin to an undercover police officer.²⁷⁵ In declining to find inducement as a matter of law, the Court noted that the undercover officer “immediately made it clear that he wanted to talk about buying large quantities of high-grade narcotics and that if [the defendant] were not interested, the conversation would end at once.”²⁷⁶ The police, in other words, provided an exculpating channel factor that an ordinary person would readily utilize to escape from inducing police conduct, and a jury could therefore reasonably infer that the defendant was not induced.

A judge should conversely weigh in the defendant’s favor the use by the police of the sorts of inculcating channel factors seen in *Connell*.²⁷⁷ The most common is the foot-in-the-door technique, which should be especially considered when the defendant pleads sentencing entrapment. The foot-in-the-door technique actually consists of two independent channel factors. First, the defendant is convinced to commit a minor offense. Second, much like a skilled salesman, the police provide him with no clear means of extricating himself. Although these appear to be the two most common inculcating channel factors, social psychologists have identified a variety of other

whether (the defendant) . . . had reasonable grounds to believe and an honest belief that he/she was in (imminent) (immediate) danger . . . you must put yourself in the position of (the defendant) . . .”; TEX. CRIM. JURY CHARGES 12:1100.120 (1999) (“[Y]ou should place yourselves in the position of the defendant at the time in question and view the circumstances from his viewpoint alone.”). The inducement element of an entrapment defense similarly requires the jury to assess how a reasonable person would respond to police pressure, and a golden-rule-type argument should be similarly permissible.

274. 356 U.S. 386 (1958).

275. *Masciale*, 356 U.S. 386 (1958).

276. *Id.* at 387.

277. *United States v. Connell*, 960 F.2d 191 (1st Cir. 1992).

means of subtly altering behavior including efforts to confuse the defendant by placing him in a situation in which ordinary expectations are contradicted²⁷⁸ and providing the defendant with a “role model” who commits the offense first and receives a benefit.²⁷⁹ Further research is needed to identify other inculcating and exculpating channel factors that may be used in sting operations.

CONCLUSION

This Note has argued that the fundamental attribution error distorts the factual inquiries involved in an entrapment defense. This persistent dispositional bias makes factfinders especially prone to attribute criminal conduct to the character traits of the defendant, and not to situational manipulations by the police. The fundamental attribution error also produces exceptionally overconfident causal attributions, which deprives the defendant of some of the benefit of the reasonable doubt. The Note has also made three suggestions for reform. First, the objective test for entrapment should be adopted as a question for the jury. Second, the jury should be instructed in a way that promotes deliberation. Finally, when evaluating entrapment claims, judges should scrutinize the sting operation for appropriate use of channel factors.

278. See discussion *supra* notes 201-202 and accompanying text.

279. See James H. Bryan & Mary Ann Test, *Models and Helping: Naturalistic Studies in Aiding Behavior*, 6 J. PERSONALITY & SOC. PSYCHOL. 400 (1967).