

A Defense of Moral Exclusion in the Law

Kenworthy Bilz¹

This article challenges a fundamental tenet of the “narrative” approach to legal judging, which I call the “moral inclusionary model,” and which is most closely associated with law and literature. This model states that before we may morally judge an offender, we must give him the opportunity to tell his story fully. Using both hard evidence from empirical psychological research in emotions and perspective-taking, and the more intuitive techniques of literary criticism, I argue that the law both does and should give equal weight to a moral exclusionary model. That is, in order to render sound, legitimate legal judgments, the law must deliberately limit the sort of mitigating stories offenders are allowed to tell—even when they are logically relevant. That is, conventionally, as both a descriptive and normative matter, we believe that evidence should lead us to particular moral evaluations. By contrast, I defend the position—again, both descriptively and normatively—that our moral evaluations often dictate what evidence is relevant in the first place.

“To understand all is to forgive all.” – *French proverb.*

In the normal run of legal cases, we think that evidence does, and should, lead us to draw moral conclusions. In this article, I argue that the reverse is often true: our moral conclusions tell us what evidence is relevant in the first place.

The first statement—that evidence→evaluations—finds descriptive support in economics and psychology, and normative support in moral philosophy and the law and literature movement. I call this view the “moral

¹ Assistant Professor, Northwestern University School of Law. J.D., University of Chicago, 1998; Ph.D., Princeton University (social psychology), 2006. Thank you to Phoebe Ellsworth, Dan Kahan, Andrew Koppelman, Janice Nadler, Claire Priest, Jeffrey Rachlinski and members of a workshop at the Center for Human Values at Princeton University and a law and psychology conference at Washington University at St. Louis, for comments. Thanks go especially to Anthony Appiah and Martha Nussbaum for giving especially helpful comments on a very preliminary draft. Their sharp skepticism at such an early stage caused me to drastically rethink my approach, hopefully for the better.

inclusionary” model of judging, and it is a compelling claim. Before you can pass judgment—say, on a criminal defendant—you have to let him tell his account in a way that would allow judges (judges, jurors) to assemble a story that is sympathetic to his circumstances. He must be allowed to invite the jurors to both understand and empathize with him. This doesn’t mean that judges oughtn’t condemn wrongdoers. It is more of a procedural point, but with the potential for substantive results: the idea is that legitimate criminal condemnation demands complete, rich, narrative trials, and ones that do not include all of the relevant particulars are undignified, illegitimate, and often even objectively incorrect.

Still, my aim is to describe and defend the flip-side, the “moral exclusionary” model; that is, that evaluations→evidence. Sometimes, it would be undignified and illegitimate, and would lead to incorrect evaluations, to let a complete, rich narrative in. Sometimes, if we are legitimately contemptuous of a criminal defendant’s unique perspective, not only is it fine that he shouldn’t be allowed to present it, but it would even be affirmatively wrong to let him do so.

To begin with a concrete example, consider the application for asylum of Amadu Bah.² Amadu is from Sierra Leone; he was a deserter from Sierra Leone’s Revolutionary United Front (RUF), a rebel army.³ He fled to the United States, and there is little doubt that had he been sent back to his home country, he would have been persecuted—probably tortured and killed—by the RUF.⁴ He should know: as a soldier, he had himself tortured and killed his fellow countrymen. He had once shot a female prisoner.⁵ On other occasions, he had used a machete to chop off the limbs and even the heads of civilians.⁶ And the United States has a sensible rule, that it will not grant asylum from persecution to people who were, themselves, persecutors.⁷ So, in deciding whether to protect Amadu, the immigration judge needed to hear two sets of relevant evidence: the likelihood that Amadu would be persecuted if he returned, and whether or not Amadu himself was a persecutor. It would be offensive to offer sanctuary from persecution to a persecutor, so being able to hear both sets of evidence is necessary to a moral, legitimate conclusion in his case.

² *Bah v. Ashcroft*, 341 F.3d 348 (5th Cir., 2003).

³ *Id.* at 349.

⁴ The Immigration Judge believed that by time Amadu’s case came up, the danger had actually passed. *Id.* at 352. However, as discussed below, even if the judge had believed he still actively faced such a risk, he would have lost his petition for asylum.

⁵ *Id.* at 350.

⁶ *Id.* at 350.

⁷ *Id.* at 351, citing 8 U.S.C. §1231(b)(3)(B)(i), which disallows the withholding of removal of an alien at risk of persecution if “the alien ordered, incited, assisted, or otherwise participated in the persecution of an individual because of the individual’s race, religion, nationality, membership on a particular social group, or political opinion”.

Amadu, however, wanted to offer a third set of evidence in his petition. He conceded that as a member of the Rebel army, he had, himself, tortured and killed. But Amadu was not just any RUF soldier; he was one of the infamous “child soldiers” of Sierra Leone. After watching the RUF light his father on fire and then rape and murder his sister, he had been offered a choice: join our cause or die.⁸ Young Amadu chose to live.

Neither the Immigration Judge nor the Fifth Circuit Court of Appeals were willing to consider this third set of evidence, stating simply that according to the governing statute, “the [persecuting] alien’s personal motivation is not relevant”⁹—even if that “motivation” was to avoid his own death. The judges interpreted the statute from the moral exclusionary model: *If* Amadu were allowed to tell his complete story, it would surely be morally relevant. But being invited to tell it in the first place itself requires moral evaluation. That is, the moral evaluation of the evidence → a decision about the relevance of the evidence at all. Congress, in this interpretation, made an *ex ante* decision that hearing persecutors’ own stories would be illegitimate—no matter how compelling they have chosen to let us listen to them. True enough, refusing to hear Amadu’s story leads us inexorably to condemn his actions—but hearing it *disables* us from doing so, and betrays his victims. There is no way to let the evidence itself render the moral conclusion. It simply is *not* the case that the moral inclusionary model is always the right one—that more evidence is better. A choice must be made *ex ante* to hear it or not, and neither choice is easy.

This example is not isolated; there are many such child soldiers, fleeing Africa and seeking asylum. In fact, the Supreme Court has recently accepted *certiorari* in order to decide how to interpret this statute, on which the Circuits have split.¹⁰ It is an extraordinarily hard question as a moral matter: does being coerced to be a persecutor obviate one’s status *as* a persecutor for purposes of asylum? But the question more generally put—when is it legitimate; indeed, ethically mandatory, to refuse to hear evidence that, if we heard it, would certainly mitigate the culpability of the person being judged?—this question is not only hard, it is common, even ubiquitous in the law.

Start with the entire Federal Rules of Evidence. “Relevant evidence,” Rule 401 states, “means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.”¹¹ The definition is self-consciously broad, and reflects the commonsense notion, reflected in the moral inclusionary model, that in order to accurately judge a

⁸ *Id.* at 349.

⁹ *Id.* at 351.

¹⁰ *Negusie v. Gonzales*, 231 Fed. Appx. 325, No. 06-60193 (5th Cir. May 15, 2007) (per curiam), *cert. granted sub nom. Negusie v. Mukasey*, No. 07-499, 2008 U.S. LEXIS 2444 (U.S. Mar. 17, 2008).

¹¹ FEDERAL RULES OF EVIDENCE 401.

situation, one must have as much information as possible. The question of relevance is the first step to admissibility: will this piece of evidence improve a jury's ultimate judgment about the merits of a case?

But no sooner do the Rules give, than they begin relentlessly taking away. Virtually every other Rule carves away what can be shown to juries, *despite* its relevance. More information, according to the Rules, is not necessarily a good thing. Some of the carve-outs at least ostensibly are designed to improve the accuracy of the jury's judgment, but many others forbid evidence for policy reasons alone, often to advance goals completely unrelated to the case at hand. This much stingier spirit is embodied by the final, catch-all exclusionary Rule 403 (on which, arguably, all of the other rules are merely a gloss): "Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence."

Because "more information is better" does comport so well with everyday notions of justice, parties frequently rebel at the notion that they should be denied the opportunity to present all of the information they can generate to make their case—particularly when the party is a criminal defendant. To condemn a person, conventional wisdom asserts, you have to know as much as possible about him and the situation he faced when he committed his acts. When we say that justice should be blind, we mean that it should blind itself to *irrelevant* particulars, not relevant ones. The bulk of evidence law doctrine is thus spent justifying the exclusions required by the Rules—a difficult task, since they so clearly subvert the strong intuition that legitimate moral judgment requires maximum information.

Nevertheless, the fact that the Rules exist at all underscore the notion that legitimate legal judging not only allows but frequently *demand*s that the "full story" *not* be told. Moreover, the Rules themselves merely embody an underlying exclusionary model that is present in the substantive law that governs it. The generosity of Rule 401 is itself something of an illusion. Even if no other rule existed, vast swaths of evidence would be excluded on Rule 401's own terms, because whether a piece of evidence is relevant in the first instance is determined by the substance of the underlying claim. For example, the defendant's lack of actual knowledge about the age of the victim in a statutory rape case may (or may not) be morally relevant, but it is *legally* irrelevant—thus evidence of his lack of *mens rea* on the issue is not admissible. The fact that the underlying substantive law declares that a man's lack of knowledge about his victim's age is irrelevant to his guilt, is an unambiguous moral conclusion that determines whether the evidence ought to be heard. The idea that the rules of evidence deliberately and justifiably hamstring parties is uncontroversial (to the lawyer, if not to the layperson), but the fact that the substantive criminal law does, and ought to do, the same thing—that doing so is both psychologically inevitable and

morally legitimate—is quite controversial, and upsetting to those committed to a liberal (in the classic, tolerant sense) system of government. This is because it is fairly easy to fool ourselves that the rules of evidence are there only to improve the “accuracy” of a verdict, not to advance a moral agenda. It’s impossible to make the same claim about the substantive criminal law: what we decide is, and is not, criminal, is very much a morally fraught endeavor.

In this article, I defend the application of the moral exclusionary model to substantive law, particularly criminal law, more generally. My argument specifically responds to a particular movement—one I do not believe to be wrongheaded in the main, but in considerable need of refinement—advocating that the rich, narrative approach of literature is, and/or ought, to be used for rendering legal judgments. Support for the descriptive branch of this movement comes from psychology research showing that decision makers in the main are not Bayesians, and are not even flawed-Bayesians. Instead, people process information by assembling it into plausible “stories,” rendering decisions according to which of a set of possible stories makes the most narrative sense.¹² The normative branch of this movement is captured by statements such as this: “This is the one vital lesson that the literary person, and hence the literary analysis of law, can uniquely teach us: she can teach us how to empathize in the hard case. It is for this teaching that the law and literature movement is singularly significant.”¹³ Support for the normative position has also come from research in psychology: inducing individuals to engage in “perspective taking” of another has, for instance, been associated with more advanced levels of moral reasoning¹⁴ and more altruism,¹⁵ and leads to lower levels of aggression¹⁶ towards and stereotyping¹⁷ of that person.

¹² See, e.g., Keith J. Holyoak & Dan Simon, *Bidirectional Reasoning in Decision Making by Constraint Satisfaction*, 128 J. EXP. PSYCH.: GEN'L 1 (1999); Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 521 (1991); Nancy Pennington & Reid Hastie, *Evidence Evaluation in Complex Decision Making*, 51 J. PERS. & SOC. PSYCH. 242 (1986); Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511 (2004).

¹³ Robin West, *Economic Man and Literary Woman: One Contrast*, 39 MERCER L. REV. 867, 874 (1988).

¹⁴ Louis Kohlberg, *Moral Stages and Moralization: The Cognitive-Developmental Approach*, in MORAL DEVELOPMENT AND BEHAVIOR 31 (T. Lickona, ed. 1976).

¹⁵ C. DANIEL BATSON, THE ALTRUISM QUESTION: TOWARD A SOCIAL-PSYCHOLOGICAL ANSWER (1991), Nancy Eisenberg & P.A. Miller, *Empathy and Prosocial Behavior*, 101 PSYCH. BULL. 91 (1987).

¹⁶ N.D. Feshbach, *Studies of Empathetic Behavior in Children*, in PROGRESS IN EXPERIMENTAL PERSONALITY RESEARCH 1 (Vol. 8) (B.A. Maher, ed. 1978); Deborah R. Richardson, Georgina S. Hammock, Stephen M. Smith, Wendi Gardner & Manuel

The moral inclusionary model's insistence on the importance of contextualization and particularization has been profoundly valuable, of course. Indeed, making us understand and sometimes even embrace the most challenging "Other" may be good literature's greatest gift, and it would be hard to deny how often the history of coercive legal actions has shown where legal decision makers could have better employed that gift. Leading advocates of the moral inclusionary model point out that both ethical literature and just law must include relevant viewpoints of all characters and citizens (especially criminal defendants). Failure to do so results in impoverished narratives that cannot result in sound judgments by either readers or legal decision makers. The emphasis on the inclusionary side of the coin is understandable—humans don't generally need to be reminded to condemn the contemptible, but they do occasionally, even often, need to have their views unsettled about what is contemptible in the first place. Taking the perspective of another is an effortful, unnatural thing to do; research reveals our default setting to be quite self-centered.¹⁸

No surprise, then, that the narrative model, particularly the moral inclusionary half of it, has been influential, spawning a growing scholarship.¹⁹ At its best, this scholarship it is not a substantive (and leaden)

Signo, *Empathy as a Cognitive Inhibitor of Interpersonal Aggression*, 20 *AGGRESSIVE BEHAV.* 275 (1994).

¹⁷ Adam D. Galinsky & Gordon B. Moskowitz, *Perspective-Taking: Decreasing Stereotype Expression, Stereotype Accessibility, and In-Group Favoritism*, 78 *J. PERS. & SOC. PSYCH.* 708 (2000).

¹⁸ Mark H. Davis, Laura Conklin, Amy Smith & Carol Luce, *Effect of Perspective Taking on the Cognitive Representation of Persons: A Merging of Self and Other*, 70 *J. PERS. & SOC. PSYCH.* 713, 721-22 (1996); M. L. Hoffman, *Interaction of Affect and Cognition in Empathy*, in *EMOTIONS, COGNITIONS, AND BEHAVIOR* 103 (C.E. Izard, J. Kagan & B. Zajonc, eds. 1984); Nancy Eisenberg, Shea, Carlo & Knight, *Empathy-Related Responding and Condition: A "Chicken and the Egg" Dilemma*, in *HANDBOOK OF MORAL BEHAVIOR AND DEVELOPMENT: VOL. 2, RESEARCH* 63 (W. Kurtines & J. Gerwitz, eds. 1991)

¹⁹ See, for example, Richard O. Lempert, *Narrative Relevance, Imagined Juries, and a Supreme Court Inspired Agenda for Jury Research*, 21 *ST. LOUIS U. PUB. L. REV.* 15 (2002); *LAW'S STORIES: NARRATIVE AND RHETORIC IN THE LAW* (Peter Brooks & Paul Gerwitz, eds., 1996); MARTHA C. NUSSBAUM, *POETIC JUSTICE: THE LITERARY IMAGINATION AND PUBLIC LIFE* (1995); ROBIN WEST, *NARRATIVE, AUTHORITY, AND THE LAW* (1993); Kathryn Abrams, *Hearing the Call of Stories*, 79 *CAL. L. REV.* 255 (1994); Jane Baron, *Resistance to Stories*, 67 *S. CAL. L. REV.* 255 (1994); Marc Fajer, *Can Two Real Men Eat Quiche Together? Storytelling, Gender-Role Stereotypes, and Legal Protection for Lesbians and Gay Men*, 46 *U. MIAMI L. REV.* 511 (1992); Daniel A. Farber & Susanna Sherry, *Telling Stories Out of School: An Essay on Legal Narratives*, 45 *STAN. L. REV.* 807 (1993); Symposium, *Legal Storytelling*, 87 *MICH. L. REV.* 2073 (1989). The movement is also recognizable in Foundation Press's popular "Law Stories" series, now twenty-four volumes strong, which tells the behind-the-law details of famous legal cases (in fields as diverse and surprising as business tax, legal ethics, and antitrust), with the ambition of supplementing the usual, more narrow pedagogy of the case method. Paul Caron,

argument that we must listen to the voices of the oppressed, or that we should read or cite more literature in judicial opinions. It is instead a methodological and procedural point about how people do, and should be invited to, process information. The model is also cautionary: if citizens cannot make narrative sense out of legal conclusions (such as verdicts), they are not likely to perceive them as either legitimate or acceptable.²⁰ If we want judges to perform their task well, then, we must offer them the full panoply of evidence that will enable them to construct a sensible story. This narrative model was summarized and explicitly endorsed by the Supreme Court, in *Old Chief v. United States*:

The "fair and legitimate weight" of conventional evidence showing individual thoughts and acts amounting to a crime reflects the fact that making a case with testimony and tangible things not only satisfies the formal definition of an offense, but tells a colorful story with descriptive richness. ... Evidence thus has force beyond any linear scheme of reasoning, and as its pieces come together a narrative gains momentum, with power not only to support conclusions but to sustain the willingness of jurors to draw the inferences, whatever they may be, necessary to reach an honest verdict.²¹

Compelling and influential as it has been, though, something must be added to the inclusionary model to make it compatible with other intuitions about justice and moral accountability. The moral inclusionary scholarship makes two large mistakes, which I try to correct in this article. The first is to urge inclusiveness towards the "Other" without reservation. I argue that such inclusiveness might be a good—but it is certainly not an unmitigated one. Moral inclusion can have deeply perverse (and empirically demonstrable) consequences, including inviting judges to trade their commitments to justice for empathy with a wrongdoer. The second mistake is failing to distinguish among different types of decision makers and institutional venues in law, implying by omission that the inclusionary model should apply everywhere and equally: to judges and juries and legislators and executives alike, and in all contexts. In fact, the moral inclusionary model delivers its strongest punch outside the courtroom, even though its advocates usually emphasize its application at trial.

To be clear, my aim is not to deny the considerable contribution that the moral inclusionary model of literature offers the law, but instead to limit its scope and define its proper place. I will defend my position not only using the techniques of ethical literary criticism itself, but by referring to research

Back to the Future: Teaching Law Through Stories, 71 U. CIN. L. REV. 405, 406-408 (2002).

²⁰ Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985).

²¹ 519 U.S. 172, 187 (1997).

in empirical social psychology about how people actually do cognitively process moral sympathies and resolve incompatible ethical narratives.

In short, while I accept the inclusionary model as far as it goes, I argue that the moral *exclusionary* model deserves equal recognition and respect. Good literature does more than enable us to empathize with characters whose moral agency we might otherwise have ignored; it also consciously (and inevitably) *selects* which characters we are to respond to in this way—and demands that we ignore others. Good law does the same: justice is not passively blind; we often must proactively blind her. Put most bluntly, at the same time great literature teaches us to empathize with the deserving, it also teaches us to despise the despicable. The legal analogue is no less blunt: sometimes, good law demands structures that will cause various decision makers to deliberately ignore structurally relevant—but morally irrelevant—points of view.

My conclusion no doubt will rankle those who think that neither law nor literature should “moralize.”²² But it is also subversive of, or at least very uncomfortable for, the mainstream in the law and literature movement, which embraces the ethical dimension of both.²³ Yet it is simply an overdue application of the lessons of the law of evidence more broadly. The exclusionary function of the law is not harsh or intolerant, any more than is the law of evidence. Instead, the moral exclusionary model provides an inescapable and even desirable opportunity to express the ethical evaluations shared (or aspired to) by a just society.

In Part I of this paper, I argue that both literature and psychology suggest reasons to be skeptical of the moral inclusionary model’s claim that full narratives simultaneously enable empathy for and, where warranted, condemnation of individuals being judged under the law. I also show that both of these disciplines offer evidence that even if we could achieve this goal, we might not want to, because the inclusionary model can lead to poor judging, in terms of both accuracy and legitimacy. In Part II, I take the lessons from literature and psychology and apply them directly to the law, suggesting the limits of the moral inclusionary model, but also describing how and when a more tempered understanding of the model (that is, adding the exclusionary model to it) can make for better judging.

²² See, for example, RICHARD POSNER, *LAW AND LITERATURE* (1998).

²³ See, for example, WAYNE C. BOOTH, *THE COMPANY WE KEEP: AN ETHICS OF FICTION* (1988); MARTHA C. NUSSBAUM, *POETIC JUSTICE* (1995); Robin West, *Economic Man and Literary Woman: One Contrast*, 39 *MERCER L. REV.* 867 (1988); Robin West, *Jurisprudence as Narrative: An Aesthetic Analysis of Modern Legal Theory*, 60 *N.Y.U. L. REV.* 145 (1985).

Part I. Moral Exclusion in Law: Lessons from Literature and Psychology

A. Literature

As I will illustrate, what I have called the moral inclusionary model has been housed within several academic disciplines—though for the most part each has seemed unaware of the other’s contributions to it. Yet its most visible and developed home has been inside the law and literature movement. This movement, of course, is multifaceted. One branch of it focuses on methods: that is, some law and literature scholars borrow the sophisticated techniques of literary criticism to suggest improvements in how the law should interpret its own texts.²⁴ Another group—not entirely distinct from the first—approaches the study of literature as a way to understand the substance and value of laws, legal structures and whole legal cultures.²⁵ At risk of oversimplifying its distinctiveness from these other branches, I direct my aim instead on what is probably the newest set of contributions to the field, which focuses on how literature’s “ways of knowing” and morally evaluating differ from how classical approaches to law do the same. This approach argues that the law can learn a great deal from literature; indeed, that in those areas where the law departs from the narrative model of knowing and judging, it is prone to normative errors.

This view of law and literature (which is tied to the larger “expressive” movement in law²⁶) is best understood as a sharp response to the law and economics movement, which has also urged the law to adopt its own vision of knowing and evaluating. Where law and economics urges aggregation, abstraction, and neutrality, law and literature urges particularism, subjectivism, and normativity. Martha Nussbaum, one of the leading advocates of the narrative approach to law, relates an encapsulated

²⁴ See, for example, INTERPRETING LAW AND LITERATURE: A HERMENEUTIC READER (Sanford Levinson & Steven Mailloux eds., 1988); STANLEY FISH, IS THERE A TEXT IN THIS CLASS? (1980), Stanley Fish, *Working on the Chain Gang: Interpretation in Law and Literature*, 60 TEX. L. REV. 551 (1982); Owen Fiss, *Objectivity and Interpretation*, 34 STAN. L. REV. 739 (1982).

²⁵ See, for example, BROOK THOMAS, CROSS EXAMINATIONS OF LAW AND LITERATURE: COOPER, HAWTHORNE, STOWE AND MELVILLE (1987); RICHARD H. WEISBERG, THE FAILURE OF THE WORD: THE PROTAGONIST AND LAWYER IN MODERN FICTION (1984); JAMES BOYD WHITE, WHEN WORDS LOSE THEIR MEANING: CONSTITUTIONS AND RECONSTITUTIONS OF LANGUAGE, CHARACTER, AND COMMUNITY (1984); JAMES BOYD WHITE, THE LEGAL IMAGINATION: STUDIES IN THE NATURE OF LEGAL THOUGHT AND EXPRESSION (1973).

²⁶ Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503 (2000); Dan M. Kahan & Martha C. Nussbaum, *Two Conceptions of Emotion in Criminal Law*, 96 COLUM. L. REV. 269, 351-53 (1996); Dan M. Kahan, *What Do Alternative Sanctions Mean?*, 63 U. CHI. L. REV. 591, 596 (1996); Kenworthy Bilz, *Self-Incrimination Doctrine is Dead; Long Live Self-Incrimination Doctrine*, forthcoming in CARDOZO L. REV. (2008); Kenworthy Bilz, *The Puzzle of Delegated Revenge*, 87 B.U. L. REV. 1059, 1088 (2007).

illustration of the difference between the two, in a passage from Charles Dickens's novel *Hard Times*. In it, a young student (Sissy Jupe) has come back from her first day at a school, and relates to a friend a question she was asked by the headmaster, who himself is a great fan of economics:

“And he said, Now, this schoolroom is a nation. And in this nation, there are fifty millions of money. Isn't this a prosperous nation? Girl number twenty, isn't this a prosperous nation, and a'n't you in a thriving state?”

“What did you say?” asked Louisa.

“Miss Louisa, I said I didn't know. I thought I couldn't know whether it was a prosperous nation or not, and whether I was in a thriving state or not, unless I knew who had got the Money and whether any of it was mine.”²⁷

Sissy Jupe—ethically engaged, demanding distributive details before she can make an evaluative call—is my literary embodiment of the moral inclusionary model.²⁸ But as appealing and compelling as her approach is, it is incomplete as a normative matter, and simply incorrect as a descriptive one, either for law *or* literature.

To see why, consider the following (admittedly exaggerated) summary of how the moral inclusionary model might evaluate Thomas Keneally's Booker Prize-winning novel *Schindler's List*.²⁹ Using the model, we would have to conclude that the novel (and the film adaptation that followed it) is morally defective. A central character, Amon Goethe, the commandant of the Plaszów labor camp, is relentlessly depicted as a “stereotyped, villainous ‘other[,],’”³⁰ whom the reader has not been invited to understand or connect with. Yet Goethe is as human as you and me, indeed, as human as Oskar Schindler and the Jews he rescues.³¹ Author Thomas Keneally has taken great pains to help us empathize with Schindler and the self-named Schindlerjuden. But when it comes to Goethe, in contrast, Keneally suffers

²⁷ CHARLES DICKENS, *HARD TIMES*, as quoted in MARTHA NUSSBAUM, *POETIC JUSTICE* 49 (1995).

²⁸ The headmaster, the evocatively-named Thomas Gradgrind, offers a bald stereotype of economic thinking. One can argue that the law and literature movement (and Charles Dickens) oversimplifies and misunderstands law and economics, *see* RICHARD A. POSNER, *LAW AND LITERATURE* 319-320 (1998), but my own objections come from another direction entirely.

²⁹ THOMAS KENEALLY, *SCHINDLER'S LIST* (1982); adapted by director Steven Spielberg, *Schindler's List* (picture citation? 1993).

³⁰ WAYNE C. BOOTH, *THE COMPANY WE KEEP: AN ETHICS OF FICTION* 203 (1988) (hereinafter *THE COMPANY WE KEEP*).

³¹ Goethe, in fact, was not just “human” in the sense that he was a fictional character whom we are told is human. As the reader most likely knows, *Schindler's List* is a true story; Goethe was a living, breathing person and the atrocities he performed, as detailed by Keneally, really happened.

from “empathic impotence,³²” for the literary ideal demands compassion “where such empathic knowledge is most difficult: of the person with the different racial heritage, the different family history, the different intelligence, or the different ambitions, goals, happiness and sorrow”³³—or, I might add, of the person with tragically misguided opinions about Jews. A good novel recognizes that “seeing others as similarly and fully human entails seeing them as individuals with their own stories to tell.”³⁴ This one did not, and because of this, Keneally’s novel was a failure; indeed, should be morally condemned. Similarly, before we could appropriately condemn Goethe in real life (say, in a court of law for his war crimes), we would need to appreciate his family history, any limits to his cognitive or emotional capacities, *everything*, in fact, that would be relevant to understanding why he did what he *did*. Condemn him, yes—but to do so with any legitimacy, we must understand him first.

The above paragraph should strike anyone, I hope, as unsettling—or even ridiculous—either as a principle of good literature or good law. Goethe was a sadist; a casual murderer whom Keneally had no obligation to paint in a sensitive way despite our sharing a common humanity with him. Literature is rife with similar examples of villains uncontroversially painted with brushes of untempered evil: *Othello*’s Iago, *Billy Budd*’s John Claggart, *Executioner’s Song*’s Gary Gilmore, and the list goes on. In each one, the author’s depiction heightens our pity for some characters (the Schindlerjuden, Desdemona, Billy Budd, Gilmore’s two murder victims) by refusing to allow sympathy for their tormentors. While not every novel does—or should—engage this technique, that some very good ones do, and do it well, is clear.

The less-exaggerated, legal analogue to the example of Amon Goethe is Amadu Bah. And it is a far harder call in Bah’s case about whether we would be morally entitled, or perhaps even compelled, to include his background as a child soldier when we are judging his particular case. But both examples demonstrate why and how moral exclusion can be both descriptively and normatively compelling. I turn now to a literary analysis (the preferred technique, of course, of law and literature) to explain in finer detail why this is so; the section that follows does the same, using research in psychology.

The moral inclusionary model demands a rich portrayal of context; a narrative fullness that essentially defines what we mean when we call something “literature,” as opposed to something that is mere description or argument. At the same time, too much context results in something equally

³² Robin West, *Economic Man and Literary Woman: One Contrast*, 39 *MERCER L. REV.* 867, 869 (1988).

³³ *Id.* at 872.

³⁴ MARTHA C. NUSSBAUM, *POETIC JUSTICE* 96 (Beacon Press 1995).

unidentifiable as literature. To some degree this is because the human mind can hold only so much detail. But this is only part of, and really the less interesting part of, the story: even were we able to process the voluminous (even endless) amount of descriptive detail a writer could deploy in generating a narrative, we wouldn't want to.

Start with what it would look like to do the opposite. This is literature that strives to achieve complete "openness;" that is, the quality of serving up to the reader "questions" rather than "answers." The ostensible goal of "open" literature is to leave everything untold, so that the reader is left acutely aware of the myriad possibilities and boundless diversity of life. In his important work *The Company We Keep: An Ethics of Fiction*, Wayne C. Booth pokes fun at this genre of literature, demonstrating that whether or not a certain measure of "openness" is desirable, a pure devotion to it is impossible. Every word and phrase selected serves to displace another, which would have had a slightly (or radically) different meaning. The author could simply continue to add "or" clauses, to "open" possibilities of interpretation. But even if the author could in this way actually arrive at, as Booth puts it, "The Great Inane," he surely "could tell no stories there."³⁵

One can make a corollary observation about the impossibility and undesirability of attempts to do the reverse: to include the complete set of particulars in a work of literature, *closing* all loopholes, instead of ensuring that all remains open. Such a work would include a total accounting of every detail and every minor player, for fear that something important might have been left out.³⁶ Martha Nussbaum, in the introduction to her *Love's Knowledge: Essays on Philosophy and Literature*, points out the weaknesses of any book written that way, when she observes that literature is valuable as a kind of filtering device for life. While "much of actual life goes by without ... heightened awareness," literature serves as "a close and careful interpretative description" which "giv[es] the reader experience that is deeper, sharper, and more precise than much of what takes place in life."³⁷ Without that filtering, a novel would be an uninteresting, unreflective chronology of the mundane details of everyday living. A reader would do better to toss such a "novel" aside and simply live his own life.

In other words, good, useful, readable literature demands that the author select which details are important, relentlessly paring away the relevant from the irrelevant. Failure to do so means no story has been told at all. We can draw an analogy here to common criticisms of economic modeling. Detractors frequently argue that economic models are too spare; that they

³⁵ THE COMPANY WE KEEP at 65.

³⁶ The stronger view of this is that one adds detail not only out of fear that important bits might be left out, but because everything actually *is* important to a full and rich telling of the story.

³⁷ Martha C. Nussbaum, *Introduction: Form and Content, Philosophy and Literature*, in *LOVE'S KNOWLEDGE: ESSAYS ON PHILOSOPHY AND LITERATURE* 3, 47-48 (1990).

leave out important variables that govern motivation and behavior. That is, the models ignore vast and important swaths of human experience, rendering them suspect as either explanations or predictors of the real world. Economists, rightly, respond that a model that is too enriched loses its value as a tool for crystallizing and clarifying all the messy variables in play. Indeed, a “model” that includes every variable that might explain the world is no a model at all; it is just “data.” The whole point of a model is to condense and weed out the irrelevant and unilluminating. In a similar way, a novel is a pared-down “model” for life, not an unfiltered replication of it.

But overspecified literature can be worse than a waste of the reader’s time; that is, worse than being simply structurally flawed and unhelpful. Drawing attention to *all* particulars has the necessary consequence of drawing our attention to at least some of the *wrong* particulars, resulting in a piece of work that gives a morally incorrect, as well as structurally unhelpful or uninteresting, account. We reject a book (or at least so much of it) that delivers what we consider to be an offensive message, or that we feel has uncritically painted a repugnant picture of the world. We expect the author to draw our attention to the *right* particulars, and ethical critics will justly hold her accountable for failures to do this.

Yet most critics who accept the ethical obligations of fiction prefer to focus on the underspecification trap. Wayne Booth provides at least one example of this kind of complaint when he explains the literary failures of the pop-fiction novel *Jaws*. He argues that the author did not—indeed, deliberately did not—humanize the victims of the primary antagonist in the book, a great white shark.³⁸ The characters are drawn as mere fodder for a sensationalist and base desire in his readers for gore; in short, the characters are fish food. The author’s failure to portray them to us as real people cons us into “enjoying the prospect of bloody death for those who don’t matter”³⁹—a sort of murderous complicity with which Booth is, and we should be, uncomfortable.

But is this really fair? Is it reasonable to expect an author to include a full fleshing out of every character who appears in her novel? The answer is of course no. As Booth hints, though does not directly say, an author need only personalize those who “matter.” When making her decision whether to paint with vivid colors or a flat gray, part of an author’s consideration is whether or not the character is necessary to drive her story along in a comprehensible manner. This is the structural sense of “matters.” Following it, we might find that by failing to flesh out certain characters, the author has left us with a confusing, cursory, or uninteresting story.

But the other sense of “matters” to which we have alluded is whether or not creating a rich sense of a particular character is *morally* necessary for the

³⁸ THE COMPANY WE KEEP at 202-204.

³⁹ *Id.* at 202-203.

story.⁴⁰ The author must, in other words, three-dimensionalize certain characters not just so that her narrative remains interesting or digestible, but also because she wants to maintain a positive relationship with the decent reader. It is this latter sense of a character's mattering which is at the heart of ethical criticism. Some perspectives simply ought to be included, not as a structural matter (to make the story interesting or comprehensible), but as a moral matter—to make the story ennobling (or at least not degrading). Full engagement with the narrative involves assessing whether the author has made good choices in this regard.

Naturally, one might disagree with the author's decision about who matters enough to warrant full inclusion; for instance, whether or not Seneca's version of *Medea* gave fair consideration to the two young boys whose murders provide foil for the main characters' (their parents) interactions, or whether Dickens's *Hard Times* was too casual in its detailing of the protagonist's long-suffering girlfriend Rachel, who, by reason of her sex, probably had more to complain about than just the labor oppression and class issues on which Dickens chose to focus. (Indeed, it was this very kind of explicitly moral choice in *Jaws* that Booth felt compelled to criticize.) But in any event, just as only some characters "matter" in the structural sense, only some characters "matter" in the ethical sense either, such that we would criticize the author if she failed to include them.

The necessity of including a fleshing out of certain characters has the obvious, but generally ignored, flip side: we must also pointedly exclude others, and for exactly the same reasons. The author must make the decision that some characters do *not* "matter" to the story. And again, one way characters might not "matter" is in the structural sense—telling their story would be stylistically confusing; including them would be superfluous and lead to a flabby tale. Exclusion on structural grounds is not especially controversial, since an author wants to create a tight, interesting narrative. But characters also might not "matter" in the moral sense; that is, including their points of view would be ethically troubling. Even if an author could write a sensible, engaging novel by including the perspectives of some characters (that is, we can think of no plausible structural grounds for exclusion), sometimes she ought to exclude them anyway, when offering their perspective would demonstrate and invite an expression of empathy toward their situation that they ought not be given. This reason for exclusion deserves much more careful attention than it has received in either ethical literary criticism, or in law and literature.

The example from *Schindler's List* illuminates both the inclusionary and exclusionary points. Much of the power of the novel (and to a much lesser extent, the movie) comes from its success in humanizing and particularizing

⁴⁰ I recognize that the two are usually related in practice. For instance, if we were to exclude a character's point of view for moral reasons, we are not likely to find that his or her inclusion nevertheless would have been structurally desirable. The concepts, however, are analytically, and importantly, distinct.

people who in less successful portrayals—and in history books—have been woodenly displayed in a tragic “lump” of “holocaust survivors.” Had Keneally failed to paint the Schindlerjuden as vividly as he did in his novel, he would have created a cartoonish, superhero melodrama about Schindler himself, limited in its ability to make us engage with the Schindlerjuden and empathize with their circumstances. The Schindlerjuden would have remained distinctly “other,” and not only would Keneally’s novel have felt incomplete and perhaps even beside the point (what is Schindler, after all, without the Schindlerjuden he saves?), but worse: Keneally’s opportunity to expand our moral capacities by recognizing and even befriending the Schindlerjuden would have been lost.⁴¹ That is, had Keneally not richly described the Schindlerjuden, he would have been vulnerable to a criticism from the inclusionary model: he would have failed to include the perspectives of characters who “matter” in at least a structural sense, and in an ethical sense as well.

Excluding the Schindlerjuden would have made for an un compelling novel, but probably not an offensive one. But consider in contrast Keneally’s decision to exclude a rich, full understanding of Goethe. We know nothing about the inner workings of Goethe’s mind; we do not know why he takes pleasure in routinely “stepp[ing] out his front door and murder[ing] a prisoner at random.”⁴² The few up-close glimpses we do get of him, far from completing and humanizing his point of view, serve to *reinforce* our notion of him as a monster, a creature whose moral core has been completely rotted out. But still, from the perspective of *structure*, Goethe is an important, even central, figure. Many of the novel’s most compelling moments focus on him. We cannot dismiss him as a minor character, whose story is tangential. As both an important character and as a human being, shouldn’t Keneally have offered him the respect of a sensitive portrayal of his circumstances?

To answer this, imagine that Keneally had realized the “error” of his ways. Before sending his manuscript off for publication, he determined that such an important (indeed, central) figure as Goethe deserved a more complete drawing, and transformed him from a caricature into a genuine character. Imagine Keneally sensitively reworking Goethe, explaining his motivations in a compelling way (his parents were rabid anti-Semites who drilled the subhumanity of Jews into his head from an early age; his schooling was drenched with similar hateful lessons which all but ensured

⁴¹ In fact, to the extent that its “cast of thousands” *did* limit our ability to become entangled more completely with a subset of the Schindlerjuden, his novel was not as effective a moral education as it could have been. I contrast it in my own mind with *The Diary of Anne Frank*, which did not as vividly display the horrors of the extermination of the Jews by Nazi Germany, yet nevertheless is a far more moving and powerful work. I attribute this to its success in completely immersing us in the thoughts and lives of Anne and her family, so that our relationship with her was far more personal and connected than with Keneally’s Schindlerjuden.

⁴² THOMAS KENEALLY, *SCHINDLER’S LIST* at 192.

his transformation into the man he became...) To the extent he was successful in truly creating a rich, thorough treatment of Goethe, we would be disabled from hating him. We might find his actions wrong, and his character, now explicable, nevertheless rotten. Still, if the portrayal is truly sensitive and convincing, we could not really *blame* Goethe for his character. Being willing to really hear his story entails slipping into his skin, and creating a relationship (even a sort of friendship, as Booth would say) that is incompatible with hating him. Of this, at the very least, we can be sure: had Keneally so written *Schindler's List*, it would be a very, very different work than it is now. More importantly, the more "comprehensive" version would be subject to harsh criticism by many decent critics. This hypothetical version of *Schindler's List* may not have committed errors of omission, but it would have committed a far larger and far worse error of commission.

Keneally didn't write that alternative novel, but others have proposed something similar. Consider the controversy surrounding CBS television's 2003 miniseries on the early life of Hitler. When the project was announced, it was immediately condemned by Jewish leaders. Their concern was that by depicting Hitler's childhood and early adulthood, the miniseries would humanize him. The head of the Anti-Defamation League stated that it was "distressing that people would spend talent, time and money to make this man human."⁴³ As reported by the *New York Times*, "Critics of the project said that whatever its intentions, a movie dealing with a young Hitler is bound to create a certain sense of sympathy... Even in a film like 'Silence of the Lambs,' the Oscar-winning performance of Anthony Hopkins as a cannibalistic killer was so captivating that audiences were alternately revolted and charmed by him."⁴⁴

In the face of the objections, CBS ended up gutting their script, ignoring most of Hitler's youngest years and focusing instead on his life after the age of 30.⁴⁵ Instead of any kind of sensitive portrayal, Hitler was depicted as "evil from day one. ... At various times in the show, Hitler is sweating, drooling, spitting food, and gushing maniacally about his passion for hatred. He is humorless and he seems to have no personal magnetism at all."⁴⁶ Thus, instead of making an ethically repugnant miniseries, executives at CBS chose instead to make a bad one. The concession did satisfy its harshest moral critics, however. The Anti-Defamation issued a statement after it aired:

"Hitler: The Rise of Evil" teaches a significant lesson of history in a powerful way. Not only does the brilliant portrayal by Robert Carlyle

⁴³ *Jewish Leader Praises Hitler Miniseries*, ASSOCIATED PRESS, May 2, 2003.

⁴⁴ Bernard Weinraub, *Planned Mini-Series on Hitler's Early Life Brings Criticism*, NEW YORK TIMES, August 20, 2002.

⁴⁵ Rob Owen, *Hitler Miniseries Better Than Average*, POST-GAZETTE (Pittsburgh, PA), May 18, 2003.

⁴⁶ Debra Neff Nathans, *'Hitler' Miniseries Too Cautious*, THE CAPITAL TIMES (Madison, Wisconsin), May 20, 2003.

show Hitler for the monster that he was, but the film shows how many times he could have been challenged and wasn't. It makes us understand how fragile democracy is and how potent evil is. ... We commend CBS and Alliance Atlantis for this important program.⁴⁷

At this point, one could object that it can *never* be morally correct to consciously marginalize another human being. That is, the critics who objected to CBS's originally-planned version of Hitler's life story are simply wrong, just as wrong as any critics of the alternate version of *Schindler's List*. While exclusion of certain otherwise-relevant points of view for stylistic reasons cannot be helped, we ought never be told that these points of view do not *matter*. If Goethe's point of view could have been included in a stylistically effective manner, it ought to have been. If a three-dimensional Hitler was necessary to a literarily- as well as morally-effective version of his life's story, then we should have been open to it. A skilled author is capable of creating a *sensitive* portrayal of evil which is nevertheless not *sympathetic*, and a good reader should be capable of evaluating the information fairly and of coming to a morally correct conclusion on her own, without the author needing to parboil the text. And *this* is the lesson that literature offers law: the moral inclusionary task might be hard, but it is mandatory.

The first problem with this objection is that it is naïve about the ability of a text, let alone a legal system, to be value neutral. The point has been made forcefully by numerous scholars: every word chosen has the author's meta-intention imbedded within it. Even the choice to use one word involves the choice *not* to use another. As explained by literary-deconstructionist-cum-law-professor Stanley Fish, "Words are intelligible only within the assumption of some context of intentional production, some already-in-place predecision as to what kind of person, with what kind of purposes, in relation to what specific goals in a particular situation, is speaking or writing."⁴⁸ These embedded intentions necessarily reveal (at least to the attentive reader) the values expressed by the writer, for which his work and sometimes even he himself may be judged.

If there's one thing the deconstructionists have made us sure of, it's that had Keneally written his alternative text, he could not have been truly neutral in his actual execution, even putting to one side the fact that the very choice to include Goethe's motivations in the first place would have been a value-laden one. (The point has been equally forcefully made, throughout the twentieth century, about the equal inability of the structures of the law to

⁴⁷ Anti-Defamation League, ADL Statement on the CBS Miniseries "Hitler: The Rise of Evil," available at www.adl.org/PresRele/HolNa_52/4258_52.asp (last visited xxxx).

⁴⁸ STANLEY FISH, *Don't Know Much About the Middle Ages: Posner on Law and Literature*, in *DOING WHAT COMES NATURALLY: CHANGE, RHETORIC, AND THE PRACTICE OF THEORY IN LITERARY AND LEGAL STUDIES* 295 (Duke University Press 1989).

be truly value-neutral.⁴⁹) It's a good guess that the words in our alternative-universe version of *Schindler's List* would have scorned Goethe's motivations, even if subtly so. To the extent they were damning at all, the text would not have succeeded in painting a genuinely sensitive picture of him. If instead his words were (equally subtly) condoning, ethical readers would be justified in criticizing the text for it.

There is one final objection to the *Schindler's List* example. It is the most important one, and the one I will address the most carefully: there is no reason we cannot be sympathetic to *both* Goethe, with his sad childhood, *and* to the victims of his adult brutality. If Goethe matters structurally as a figure central to the telling of Keneally's story, and he also matters in the moral sense just by virtue of his being a human, then there is no ethical problem with sympathizing with his plight. His viewpoint should be included, with any error in the telling in favor of sympathy, in honor of our common humanity. To use the more compelling, and tougher, case of Amadu Bah, the judge in his case—and in all cases like it—*must* invite himself to sympathize with Bah by hearing and weighing his complete story. He is the central figure in his asylum hearing; his plight deserves sympathy; and so his entire narrative must be considered before he can be legitimately, legally judged. We can empathize with Bah-the-child-soldier, as well as with the victims of Bah-the-monstrous-revolutionary. This the dual-compassion argument.

In summary, the dual-compassion argument misfires because it fails to recognize that our attentions and sympathies are a scarce commodity. The point is not merely quantitative; that is, the problem isn't just that we don't have the time or mental capacity to feel sorry for everyone (although this is certainly true, and a real problem for the inclusionary model). More fundamentally, alignment with one set of interests *precludes* our ability to align with another set, because they may be, as is the case with Goethe and the Schindlerjuden, simply incompatible. Why? Seeing someone as having been wronged necessitates the recognition that they have been wronged by *someone*, or more precisely, by someone responsible. Victimization invites—indeed, fully depends on—blame. Otherwise, the harm suffered is just bad luck, and we might pity the person who has been hurt, but to feel that he has been wronged would be strange. As the philosopher Jean Hampton said, “However much we may sorrow over our bad fortune, when some damage is threatened or produced by natural forces or by accidents, we

⁴⁹ See, for example, Owen Fiss, *Objectivity and Interpretation*, 34 STAN. L. REV. 739 (1982); Robert Hale, *Coercion and Distribution in a Supposedly Noncoercive State*, 38 POL. SCI. Q. 470 (1923); Karl Llewellyn, *Some Realism About Realism—Responding to Dean Pound*, 44 HARV. L. REV. 1222 (1931); Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 88 HARV. L. REV. 1685 (1976). But see Herbert Wechsler, *Toward Neutral Principles of Constitutional Law*, 73 HARV. L. REV. 1 (1959).

do not experience that special anger that comes from having been *insulted*.”⁵⁰

When one person wrongs another, in addition to inflicting any objective harm on the victim, there has been a message sent, which says “I count but you do not,’ ‘I can use you for my purposes,’ or ‘I am here up high and you are there down below.’”⁵¹ If we were to feel sorry for the wrongdoer, that itself adds its own nasty message to the one the wrongdoer himself sent.⁵² This lesson is learned early in human development. Imagine a school yard bully shoves another child, and the teacher mediating the dispute informs the injured child that the bully “didn’t really mean it.” If the bully really didn’t mean it, then there was no insult, and the child knows that there is no reason at all to be angry.⁵³ But what if the bully did mean it, and the teacher instead explains, “but he just can’t help picking on other children,” then the child has learned something similar. Outrage at the aggressor is a way of recognizing the victim *as* a victim. Being angry with the bully engages an alliance with the victim; being an apologist betrays an alliance with the aggressor.⁵⁴

To the extent that we are disabled from seeing Goethe as the *culpable* cause of the suffering of the inmates of the Plaszów camp, we are disabled from seeing them as having been wronged—with that special empathy that comes with such recognition—at all. Simply finding some other party to blame is no solution. Should we decide that higher ranking officials in the Third Reich really is to blame for what happened at Plaszów, then the moral inclusionary model would demand their perspective be fully considered before judging them. Or should, perhaps, Goethe’s parents really be considered culpable, because they raised a monster like Goethe, then the concerned reader/judger is entitled to a sensitive (and potentially exonerating) portrayal of *them*.⁵⁵ And so on. After all the dominoes have

⁵⁰ Jean Hampton, *Forgiveness, resentment and hatred*, in FORGIVENESS AND MERCY 35, 43-44 (Jeffrie G. Murphy and Jean Hampton, 1988). See also FRITZ HEIDER, *THE PSYCHOLOGY OF INTERPERSONAL RELATIONS* (1958).

⁵¹ Jeffrie G. Murphy, *Forgiveness and resentment*, in FORGIVENESS AND MERCY 14, 25 (Jeffrie G. Murphy and Jean Hampton, 1988).

⁵² Kenworthy Bilz, *The Puzzle of Delegated Revenge*, 87 B.U. L. REV. 1059 (2007).

⁵³ Piaget showed that children understand the meaning, and the import, of intent, very early in their development. JEAN PIAGET, *THE MORAL JUDGMENT OF THE CHILD* (1932). For a review on the voluminous experimental literature demonstrating the developmental trajectory of moral sensibilities, see John M. Darley & Thomas R. Shultz, *Moral Rules: Their Content and Acquisition*, 41 ANN. REV. PSYCH. 525 (1990).

⁵⁴ Kenworthy Bilz, *The Puzzle of Delegated Revenge*, 87 B.U. L. REV. 1059 (2007).

⁵⁵ Assuming, of course, that the inclusion would make stylistic sense. Omission of points of view for structural reasons is not controversial. Yet I should think that

fallen, and everyone is revealed to be innocent because himself a victim, we come to one of two conclusions. Either no one is to blame (or maybe it is “God’s fault”)—a result unsatisfying to anyone with a respect for individuals’ moral agency—or, everyone is to blame (it’s “society’s fault”). Since, however, everyone is a part of society and so equally both victim and aggressor, victims are as much to blame for their plight as anyone else. Under this analysis, then, the Plaszów inmates were just as capable (or more accurately, incapable) of stopping the violence against themselves as was Goethe. Under either the no-one-is-to-blame view, or the everyone-is-to-blame view, we are disabled from the particular type of compassion for the camp inmates that goes with identifying them as victims of a blameworthy wrong. And again, the ethical reader would feel rightly indignant about such a bottom line.

Some literature, true enough, is designed to make the wrongdoer the central figure of the narrative as a “portrait of evil.” It is designed to show the ways in which a soul can become corrupted, and sometimes—unlike the producers of “Hitler: The Rise of Evil”—the authors manage to do this job quite well. Instead of making us empathize with the “protagonist,” however, the point of this kind of literature seems to be merely to give the reader the perverse thrill of seeing evil up close, as in the case of Bret Easton Ellis’s *American Psycho* or Norman Mailer’s *The Executioner’s Song*, both of which detail the lives of serial killers (the first fiction, the second deeply fictionalized). But this type of depiction is unhelpful to the dual compassion argument, as their subjects are not detailed in a way that makes us humanize and understand their personalities. Their stories are structurally full, but not morally full: we are not invited to sympathize. On the contrary, the point of the exercise, as appreciated by the Anti-Defamation League in evaluating CBS’s latter, lesser miniseries, seems to be to further alienate the reader from the subject, in much the same way that looking at a bug under a magnifying glass serves to enhance, not diminish, our horror.

Equally unhelpful are the type of novels in which no one is a wrongdoer, but everyone seems to be a “victim” of fate, or circumstances, or the frailties of the human condition. Kurt Vonnegut, Jr. is a master of this species of novel; a talent he acknowledged in his preface to *Slaughterhouse-Five* by noting his own father’s accusation that “[y]ou never wrote a story with a villain in it.”⁵⁶ But again, these novels don’t prove the point that we can feel acutely sorry for both victims and wrongdoers. Again, Vonnegut’s characters do not really seem like victims at all, since without someone to blame, we never have a sense of outrage at the misadventures that befall them. If anything, this sort of novel gives evidence for the point that the human psyche is incapable of feeling sharply that a person has been wronged

blaming Goethe’s parents for the horrors of the holocaust would stylistically, as well as morally, necessitate an explanation.

⁵⁶ KURT VONNEGUT, JR., *SLAUGHTERHOUSE-FIVE* 8 (Dell Publishing, 63rd prtg., 1991).

if it is presented with the assumption that *everyone* has been wronged. Pity, or exasperation, yes; but indignation and deep compassion, no.

There *is* a type of novel which at first glance appears like it could validate the dual-compassion argument. These novels are a cross between the above two types, in that they genuinely attempt to show a wrongdoer who is depraved, yet still worthy of our sympathy. (The first CBS version of Hitler's life story aspired to be this; we can't know how it would have turned out.) These kinds of novels require impressive literary chops on the part of the author, as their ambition is to make us simultaneously feel contempt and empathy towards the primary protagonist. This list might include Emily Bronte's *Wuthering Heights* (Heathcliff), Patricia Highsmith's *The Talented Mr. Ripley* (Tom Ripley), Vladimir Nabokov's *Lolita* (Humbert Humbert), and Shakespeare's *MacBeth* (Lady MacBeth).

To preface the application of this type of novel to the law, the moral inclusionary model would invite a comparison to the defendant in a criminal trial. The defendant is unambiguously the "protagonist" of the criminal trial, accused of behaving in a deeply anti-social fashion, and faced with the judgment of, prototypically but obviously not always, a jury of his peers (whose analogue are the readers). How should the trial be "written"? The moral inclusionary model would suggest that for a verdict to be legitimate—or in Justice Souter's words, "honest"⁵⁷—the presentation must allow those who would pass judgment to fairly understand the defendant, even though it might also condemn him. Think again of Amadu Bah: the moral inclusionary model would say that the fact that he was a child soldier would unambiguously be relevant, but the judge would be free to still deny him asylum on the grounds of his own persecutory acts.

One of the most sincere, self-conscious examples of this kind of novel is Richard Wright's *Native Son*.⁵⁸ In it, a young man named Bigger Thomas kills two women: the daughter of his employer in what seems to be an accident; and his girlfriend, in cold blood. The novel focuses on Bigger's abduction and trial for these crimes, and along the way details the indignities of his life in Chicago's appallingly segregated South Side ghetto. The reader is invited to empathize with his circumstances and understand the forces that made him into the twisted human he has become. Wright desired to make Bigger despicable, but nevertheless explicable given his background as a victim of shattering poverty and unrelenting racism: "I didn't want to kill",

⁵⁷ *Old Chief v. United States*, 519 U.S. 172, 187 (1997)

⁵⁸ RICHARD WRIGHT, *NATIVE SON* (1940). I do not mean to imply that Wright tried to hamhandedly produce some sort of blunt, moral cautionary tale to White America—the book would not be worthy of its status as a great work of literature if so. For the complex motivations and ambitions of Wright in producing the novel, see his essay, "How 'Bigger' Was Born," in the 1998 Harper-Collins edition of the novel.

Bigger shouted. "But what I killed for, I am! It must've been pretty deep in me to make me kill."⁵⁹

The issue is not, of course, whether these sorts of novels can be attempted, it is whether they can succeed in a way that holds lessons for the law. Did Wright—a careful, intelligent, talented writer—achieve his goal? He does succeed in making us feel ambivalent, at the very least, toward Bigger. Bigger is a monster, yes, but he is also a pathetic creature who has been twisted from having suffered many wrongs himself. But Wright's task cannot be so easy as making us feel sorry for Bigger in one instance and angry at him in another. We must feel these emotions at the same time. Ultimately, Wright achieves the confusion he is seeking in his readers only through a bit of sleight-of-hand: we feel sorry for Bigger in his first "murder," which Wright paints as accidental. Quite apart from that, we despise him for his second, passionless murder of Bessie, his girlfriend.

Wright makes it crystal clear that Bigger did not intend to murder Mary, his white employer's daughter. He had covered her mouth with a pillow to stifle her drunken articulations, so that her blind mother who had just entered the bedroom would not hear them. Had he been discovered in her room (and he was only there because she was incapable of getting there herself, as drunk as she was), all hell would have broken loose. In fact, Bigger does not even discover what he has done until he checks to see if she is all right after her mother leaves.

Panicked, Bigger covers up his crime by disposing of Mary's body in a clumsy way, and allows Mary's boyfriend to take the blame for her disappearance. His behavior after Mary's death is far from laudable, but throughout, the reader never feels as though Bigger is *contemptible* for these acts. In fact, while the reader feels sorry that Mary is dead, the reader does not feel sharply that she has been *wronged*. Not only was her death an accident, but she and her foolishly idealistic boyfriend Jan had, after all, forced Bigger to come along in their late-night alcoholic escapade—as a brand-new, black servant, Bigger was in no position to disobey their orders. Her drunkenness put Bigger in the even more impossible position of having to cart her home and put her into bed after Jan disappeared. Bigger was not being unreasonable in his fear of the consequences should Mary's mother have discovered him in her (inebriated) daughter's bedroom, and stifling her noises with a pillow seems at worst a very careless act by a panicked man in dangerous situation. Mary's death seems as much her own fault as Bigger's—perhaps even more so. The reader's empathy with Bigger at this point in the novel remains untainted by the sense that he has committed a grave wrong, or that either Jan or Mary have been *victimized*; what happened up to this point was all a tragic accident.

Perhaps realizing this, Wright then tips the balance so that the reader will not leave the novel with the feeling that Bigger has merely been the

⁵⁹ RICHARD WRIGHT, *NATIVE SON* 429 (1940) (Harper-Collins edition 1998).

unfortunate victim of vicious circumstance. His goal, again, is to make us sympathize with Bigger *despite* being outraged by him. Once the “truth” of Mary’s demise comes out (the real truth, of course, we know would never have been believed), Bigger flees. While on the lam, he brutally rapes and murders his black girlfriend, Bessie, because he feels that she will be a liability to him in his escape, and may even give him away. At this point, the reader is presented with a perfectly unredeeming, calculating Bigger, who murders in a spectacularly gory and painful way, for a vanishingly small gain to himself (with the rape thrown in seemingly for not other reason than to underscore the senselessness of the crime). We are outraged and disgusted at Bigger now, and our sympathy for Bessie is complete and unmitigated by any sense of just deserts, unlike with Mary and Jan.

It is only at this point that our confusion gels. Bigger is an object of our sympathy for the first death; he is thoroughly contemptible for the second. Yet just because we feel two ways about Bigger does not prove the argument Wright set out to demonstrate, or that the moral inclusionists urge. We might feel sorry for a man who is murdered, but *not* sorry for him for murdering his own wife the night before. Wright has not proved that one can feel sorry for both the victim and wrongdoer at the same time, in reference to the same crime, and it’s not clear that anyone can. If anything, *Native Son* undermines the dual-sympathy point, because it shows that we were disabled from feeling real anger toward Bigger until his second murder. Perhaps partly because he realized that he had been too successful in making us hate Bigger for what he did to Bessie, in the extended trial scenes at the end of the novel, Wright barely gives her a mention.⁶⁰ By once again shifting our picture of the victim, we are left unsettled in our sympathies.

But this unsettling is more mirage than truth. To the extent he is successful in making us hate Bigger, we feel sorry for his victims. But when Wright instead makes us feel sorry for Bigger, we are no longer angry at him and even feel anger at his “victims” for putting him in such a position.⁶¹ Even as masterful and ambitious a work as *Native Son* does not manage to make the reader feel both pity and outrage towards the same person at the same time. If it is this hard to do in literature, can it really be a lesson we should demand the law adopt in delivering its own moral judgments? But before I try to answer that, the next section suggests even more problems for the moral inclusionary view.

⁶⁰ Of course another reason Wright leaves out much mention of Bessie is to reflect the social reality that an impoverished, black rape-murder victim simply will not get the criminal justice system’s attention in the way a wealthy, white victim will. Cf. DAVID C. BALDUS, CHARLES A. PULASKI, AND GEORGE WOODWORTH, EQUAL JUSTICE AND THE DEATH PENALTY (1990).

⁶¹ I am not the first to note that Wright’s novel exploits emotional vacillation to make its point. As Harold Bloom wrote, “Either Bigger Thomas is a responsible consciousness, and so profoundly culpable, or else on the white world is responsible and culpable.... Wright ... was not able to choose.” *Introduction*, in RICHARD WRIGHT, *NATIVE SON* 1 (H. Bloom ed. 1988).

B. Psychology

Succeeding at the dual-compassion task is not just difficult as a literary matter, but as a psychological matter as well. The ambition of the moral inclusionary model is not, per se, to make us condemn a wrongdoer or his wrongdoing less—unless of course a fuller appreciation of him and his circumstances makes such mitigation warranted—but instead to give him, and ourselves, the dignity of allowing him to tell his story in whole. The value of allowing him to paint a complete picture is perhaps that we might “hate the sin, love the sinner,” or perhaps more aptly, “Hate what is detestable in the sinner, love or at least tolerate the rest.” Or, avoiding such religious overtones, the argument is simply that it is illegitimate to condemn someone without understanding him fully; that is, without allowing him to give a full accounting of how he came to have done what he did—and this mandate applies with full force, perhaps even more strongly, in the legal setting.

The question for the moral inclusionary model, then, is whether and in what circumstances it is actually possible to both fully condemn and fully comprehend. The moral inclusionary thesis can be divided into two separate claims: one, that doing so is possible; and two, that it is desirable.

The first claim finds some support in psychological research on attitudes. Attitudes describe our positive and negative evaluations of various objects, persons, things, situations and the like, which are known as “attitude objects”—I focus only on persons as attitude objects here. Attitudes are comprised of, and explained by, underlying emotions (how I feel towards the person), cognitions (what I believe to be true about a person), and even behaviors (how I behave toward him or her). The study of attitudes, both how they are structured and how they change, is enormously complex, and has generated a spectacular amount of research in social psychology.⁶² For purposes of the moral inclusionary model, however, the most relevant subset of this research concerns *ambivalent attitudes*.

In theory, you could add up all of all of your cognitions, emotions and behaviors toward a person and measure it as a composite “attitude” on a unipolar scale from positive to negative. In other words, when considering it as a whole, you would feel positively toward a person, or negatively toward him, or something in between.⁶³ Indeed, attitudes are usually measured in this way, as, for instance, when pollsters have respondents indicate where they would place themselves on a “feeling thermometer” toward a certain

⁶² For excellent reviews, see Alice H. Eagly & Shelly Chaiken, *Attitude Structure and Function*, in *THE HANDBOOK OF SOCIAL PSYCHOLOGY*, 4th ed. 269 (D.T. Gilbert, S.T. Fiske & G. Linzey, eds. 1998) and Richard E. Petty & Duane T. Wegener, *Attitude Change: Multiple Roles for Persuasion Variables*, in *THE HANDBOOK OF SOCIAL PSYCHOLOGY*, 4th ed. 323 (D.T. Gilbert, S.T. Fiske & G. Linzey, eds. 1998).

⁶³ Klaus Jonas, Philip Broemer & Michael Diehl, *Attitudinal Ambivalence*, 11 *EUR. REV. SOC. PSYCHOL.* 35, xx (2000).

politician. But what does it mean when the summary evaluation about a person is neither hot nor cold, but lukewarm? Interpreting that attitude is ambiguous. It could mean the perceiver is neutral or even indifferent; but, it could also mean that she “feel two ways about” him. This is the precise psychological meaning of ambivalence: one might dislike a person intensely on one dimension, but simultaneously feel some affection for her on another, and vice versa.⁶⁴ Ambivalence can occur within or across different attitudinal classes: I might have positively valenced *emotions* toward a person, but negatively valenced *cognitions* (as an example, loving a partner despite knowing that she is unfaithful). Or, I might have *both* negatively and positively valenced cognitions: I believe my spouse is faithful, but also believe that he is not a good parent. Or, I might have both negatively and positively valenced emotions toward the same person: I love a friend, but am also embarrassed by him.⁶⁵

Ambivalent attitudes toward entire social groups are ubiquitous, even universal.⁶⁶ And of course, we also frequently feel ambivalent about single, particular persons. Still, while ambivalent attitudes are common, normally, they stay under the radar of a person’s consciousness.⁶⁷ That is, even where

⁶⁴ Jonas, Broemer & Diehl, *Attitudinal Ambivalence* at xx.

⁶⁵ Jonas, Broemer & Diehl, *Attitudinal Ambivalence* at 41-42.

⁶⁶ Susan T. Fiske & Amy J. C. Cuddy, *Stereotype Content and Relative Group Status Across Cultures*, in *SOCIAL COMPARISON PROCESSES AND LEVELS OF ANALYSIS: UNDERSTANDING CULTURE, INTERGROUP RELATIONS AND COGNITION* 249 (S. Guimond, ed., 2006). For instance, people often have ambivalent attitudes toward particular racial or ethnic groups. Irwin Katz & R. Glen Hass, *Racial Ambivalence and American Value Conflict: Correlational and Priming Studies of Dual Cognitive Structures*, 55 *Attitudes & Cognition* 893 (1988). Men often have well-documented ambivalent attitudes toward women. Peter Glick & Susan T. Fiske, *The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism*, 70 *J. PERS. & SOC. PSYCH.* 491 (1996). People generally have ambivalent attitudes toward the elderly, Amy J. C. Cuddy & Susan T. Fiske, *Doddering, but Dear: Process, Content, and Function in Stereotyping of Older Persons*, in *AGEISM: STEREOTYPING AND PREJUDICE AGAINST OLDER PERSONS* 3 (T. Nelson, ed., 2002); working mothers, Amy J. C. Cuddy, *When Professionals Become Mothers, Warmth Doesn’t Cut the Ice*, 60 *J. SOC. ISSUES* 701 (2004); the handicapped, KATZ, SODER, and others. Some individuals, however, are more likely to feel ambivalent attitudes more strongly and more often than others, however. Megan M. Thompson & Mark P. Zanna, *The Conflicted Individual: Personality-Based and Domain-Specific Antecedents of Ambivalent Social Attitudes*, 63 *J. PERS.* 259 (1995). And, some may be more bothered by the feelings than others. Ian R. Newby-Clark, Ian McGregor & Mark P. Zanna, *Thinking and Caring About Cognitive Inconsistency: When and for Whom Does Attitudinal Ambivalence Feel Uncomfortable?*, 82 *J. PERS. & SOC. PSYCH.* 157 (2002).

⁶⁷ It is still possible to measure ambivalent attitudes that are below the surface of awareness. Bertram Gawronski & Galen Bodenhausen, *What Do We Know About Implicit Attitude Measures and What Do We Have to Learn?*, in *IMPLICIT MEASURES OF ATTITUDES* 265 (B. Wittenbrink & N. Schwarz, eds., 2007). One such technique to measure them that has received considerable attention is the Implicit Association

people hold beliefs that are evaluatively inconsistent with one another, the inconsistency is of little consequence because they remain unaware of it unless explicitly confronted with an event or situation (such as the need to make a decision) that brings the inconsistency to the forefront.

In the context of the law, though, the moral inclusionary model would purposely make attitudinal ambivalence salient. The model invites, and even urges, us to be open to the possibility of experiencing ambivalence in our attitudes toward anyone we legally judge. It tells us that we can and ought to feel both outrage about the wrong, and empathy for the wrongdoer. Those passing judgment *should* be confronted with the panoply of relevant information about a wrongdoer, inviting compassion “where such empathic knowledge is most difficult.”⁶⁸ It is only by “seeing them as individuals with their own stories to tell” that we can see the wrongdoer as “similarly and fully human.”⁶⁹ This would entail “experienced” ambivalence, where the person holding evaluative attitude is subjectively aware of the conflicts.⁷⁰

Despite the evidence about ambivalent attitudes generally, which supports the moral inclusionary model, whether or not ambivalence is possible in the way the model would like is not a simple issue. For one thing, advocates of moral inclusion have not been particularly clear about just what sort of ambivalence they are inviting. There are several possibilities, which can be organized across the response classes already described: cognitive and emotional.⁷¹ The least problematic form of

Test, which measures the speed with which one associates groups (say, women versus men) with particular positive versus negative terms. *See generally* Kristin A. Lane, Mahzarin R. Banaji, Brian A. Nosek, & Anthony G. Greenwald, *Understanding and Using the Implicit Association Test: IV—What We Know (So Far) About the Method*, in *IMPLICIT MEASURES OF ATTITUDES 59* (B. Wittenbrink & N. Schwarz, eds., 2007). Though the techniques for measuring implicit attitudes, and opinions about what, exactly, they are measuring, remain somewhat controversial; *see generally* Gregory Mitchell & Philip E. Tetlock, *Antidiscrimination Law and the Perils of Mindreading*, 67 *OHIO ST. L. J.* 1023 (2006) and Brian A. Nosek, Anthony G. Greenwald & Mahzarin R. Banaji, *The Implicit Association Test at Age 7: A Methodological and Conceptual Review*, in *AUTOMATIC PROCESSES IN SOCIAL THINKING AND BEHAVIOR* 265 (J. A. Bargh, ed., 2007); the existence of implicit attitudes themselves is not.

⁶⁸ West, *Economic Man and Literary Woman* at 872

⁶⁹ NUSSBAUM, *POETIC JUSTICE* at 96.

⁷⁰ Jonas, Broemer & Diehl, *Attitudinal Ambivalence* at 55.

⁷¹ I am skipping the third category for now: behavior. It is true that behaviors toward an attitude-object can comprise part of one’s composite attitude, but in the context of the moral inclusionary model, normally one does not consider how one has behaved in the past toward a person being judged in assessing how one feels toward the person as a whole. (That this does happen, though, is the main insight of Bem’s “self-perception theory,” which shows that sometimes we look to our behaviors to assess our attitudes, because our cognitions might be opaque even to ourselves. Daryl Bem, *Self-Perception Theory*, in *ADVANCES IN EXPERIMENTAL PSYCHOLOGY*, VOL. 6, 1 (L. Berkowitz, ed., 1972).) True enough, behavior comes into play with the moral inclusionary model—but as a *response* to the attitude, not, primarily, a

ambivalence is purely cognitive. You might know two competing facts about a person, but depending on the relative magnitude of importance of those facts, the ambivalence is easy enough to solve into a single decision.⁷² For instance, the burglar who carefully and politely picks the lock is still a blameworthy thief, even if we evaluate him slightly more positively than the burglar who kicks the door off its hinges to gain entry. Our knowledge that he is polite enough to minimize gratuitous damage is inconsistent with our knowledge that he is a thief, but since the latter is more important than the former, the summation of the inconsistent cognitions into a unipolar attitude is easy enough to do.

The goal of the inclusionary model, though, seems to be that we invite something different: a form of *emotional* ambivalence. The idea is not just to make judges aware of the inconsistent “facts” they must process in rendering a legal evaluation of a wrongdoer (that he hurt the victim, but also, say, that he himself was abused as a child), but to translate those cognitions into *sympathy* coupled with *outrage* towards him. Purely emotional ambivalence is far more elusive than the purely cognitive variety. Many social psychological researchers do not believe that inconsistent emotions can be held simultaneously at all.⁷³ Even those who do accept the possibility do not believe that *all* emotions can be bi-dimensional—some truly are mutually exclusive, such as calm/tense.⁷⁴ The evidence that does support the possibility of bi-dimensional emotions has focused only on the happy/sad pair; no one has examined whether sympathy/outrage are another exception to the usual mutual-exclusivity rule. Moreover, even in the rare instances where conflicting emotions can be active at the same time in the same person, the result is “a poor guide for behavior and is therefore posited to be unpleasant, unstable, and often short-lived.”⁷⁵ Not as elusive, but more complex, are ambivalent attitudes where cognitions conflict with emotions. This is the sensation of one’s “heart and mind” pulling in different directions. This, too, might be what the moral inclusionary model urges:

component of it. That is, judges must render a decision—such as a verdict or a sentence.

⁷² This process is described in part by Ajzen’s classic Expectancy-Value Model, which describes attitudes as a composite function of the various subjective values for the attributes a person associates with an attitude object, multiplied by the strength of those associations. Icek Ajzen, *Nature and Operation of Attitudes*, 52 ANN. REV. PSYCH. 27, 30-32 (2001).

⁷³ James A. Russell, *Core Affect and the Psychological Construction of Emotion*, 110 PSYCH. REV. 145, 148 (2003); James A. Russell & J. M. Carroll, *On the Bipolarity of Positive and Negative Affect*, 125 PSYCH. BULL. 3 (1999).

⁷⁴ John T. Cacioppo, John T. Larsen & Peter McGraw, *Can People Feel Happy and Sad at the Same Time?*, 81 J. PERS. & SOC. PSYCH. 684, 692-695 (2001).

⁷⁵ Cacioppo, Larsen & McGraw, *Can People Feel Happy and Sad at the Same Time?* at 692. See also John T. Cacioppo & Gary G. Berntson, *Relationship Between Attitudes and Evaluative Space: A Critical Review, With Emphasis on the Separability of Positive and Negative Substrates*, 115 PSYCH. BULL. 401, 403 (1994).

knowing that someone engaged in sanctionable behavior, yet nevertheless feeling empathy for the reasons he did it. However, as with purely emotional ambivalence, this kind of ambivalence tends to feel unpleasant to the one actively experiencing it,⁷⁶ and leads to attitudes that are not particularly stable over time.⁷⁷

But perhaps this is good enough. If a judge can be made to appreciate the complexity of moral evaluation by experiencing vacillating emotions, even if only fleetingly, then the first part of the moral inclusionary thesis (that people *can* simultaneously condemn and sympathize) is at least weakly supported by psychology, just as the rapidly vacillating emotions induced by reading *Native Son* at least weakly supports the descriptive part of the model. Yet there is another claim to be dealt with, and that is the notion that if a judge can be made to empathize in this way with a wrongdoer, he *ought* to be.

The moral inclusionary model must first address the problem that only some of the properties of experienced ambivalence are good. Attitudinal ambivalence, as already mentioned, is unpleasant (“aversive”), and as such people are motivated to reduce it. One way they attempt to do so seems at first cut like a very strong point in favor of the moral inclusionary model: they carefully attend to the inconsistent information, expending more cognitive resources to make their ultimate judgment than they otherwise would.⁷⁸ Better still, this increased attention leads them to be more persuaded by strong arguments and more impervious to weak ones.⁷⁹

Unfortunately, other features of ambivalent attitudes are far more troublesome for the moral inclusionary model. For one, ambivalent attitudes often become polarized in their evaluation of the person, resulting in

⁷⁶ LEON FESTINGER, *A THEORY OF COGNITIVE DISSONANCE* (1957); FRITZ HEIDER, *THE PSYCHOLOGY OF INTERPERSONAL RELATIONS* (1958). Some have argued that it is not the dissonance itself that is aversive, but the implications of it (such as to one’s self-concept). See, for example, Claude M. Steele, *The Psychology of Self-Affirmation: Sustaining the Integrity of the Self*, in *ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY* 261 (L. Berkowitz, ed., 1988). For my purposes, however, this is a distinction without a difference, since for whatever the reason, the experience of dissonance motivates the one experiencing it to reduce it.

⁷⁷ Jonas, Broemer & Diehl, *Attitudinal Ambivalence* at 39, 64; Klaus Jonas, Philip Broemer & Michael Diehl, *Attitudinal Ambivalence as a Moderator of the Consistency Between Attitudes and Behaviors*, 31 *ZIETSCHRIFT FÜR SOZIALPSYCHOLOGIE* (2000)

⁷⁸ Philip Broemer, Klaus Jonas & Michael Diehl, *Effects of Evaluative (In-)Consistency and Perceived Self-Efficacy on Attitude Formation*, *SWISS J. OF PSYCH.* (2003); Gregory R. Maio, David W. Bell, & Victoria M. Esses, *Ambivalence and Persuasion: The Processing of Messages About Immigrant Groups*, 32 *J. EXP. SOC. PSYCH.* 513 (1996).

⁷⁹ Gregory R. Maio, David W. Bell, & Victoria M. Esses, *Ambivalence and Persuasion: The Processing of Messages About Immigrant Groups*, 32 *J. EXP. SOC. PSYCH.* 513 (1996).

exaggerated positive attitudes if exposed to positive information about the person, and exaggerative negative attitudes when exposed to negative information.⁸⁰ Worse, more than just their attitudes may get exaggerated. Ambivalent attitudes may also trigger an impulse to engage in extreme behavior, either positive or negative, toward the object of ambivalence (which here, of course, is the individual being judged). This phenomenon, known as “response amplification,” is a robust finding for how people respond to ambivalence—namely, they behaviorally buttress either the positive information they have about the target, or the negative, in order to reduce their sensation of ambivalence.⁸¹ (Those familiar with the classic phenomenon of cognitive dissonance reduction in psychology will recognize this move.)⁸² Another way people attempt to reduce ambivalence is to turn to their peers to decide how to view the attitude object.⁸³ Again, the desire to eliminate unpleasant inconsistency can cause people to turn to methods not governed by the merits of the issue, but by ease and expediency. Advocates of the moral inclusionary model would not take much comfort in the idea that people resolve complex, mixed feelings about a wrongdoer by simply recalibrating their attitudes to comport with those of their peers.

Even if the ambivalence could be actively maintained instead of quickly collapsing into either a positive or negative view of the target, judges aren’t asked to form an *attitude* about the person being judged, they are asked to engage in a *behavior*; namely, making a decision in the form of rendering a verdict, delivering a sentence, or fixing damages. The model implicitly relies on the belief that complexity in judges’ attitudes will lead to higher-quality (or at least not lower-

⁸⁰ R. Glen Hass, Irwin Katz, Nina Rizzo, Joan Bailey, & Donna Eisenstadt, *Cross-Racial Appraisal as Related to Attitude Ambivalence and Cognitive Complexity*, 17 PERS. & SOC. PSYCH. BULL. 83 (1991); Patricia W. Linville & Edward E. Jones, *Polarized Appraisals of Out-Group Members*, 38 J. PERS. & SOC. PSYCH. 689 (1980).

⁸¹ This is known as “response amplification.” David Bell & Victoria M. Esses, *Ambivalence and Response Amplification: A Motivational Perspective*, 28 PERS. & SOC. PSYCH. BULL. 1143 (2002); Katz & Hass, *Racial Ambivalence and American Value Conflict* at 899; Irwin Katz, Joyce Wackenhut & R. Glen Hass, *Racial Ambivalence, Value Duality, and Behavior*, in PREJUDICE, DISCRIMINATION, AND RACISM 35 (J. F. Dovidio & S. L. Gaertner, eds., 1986).

⁸² For the classic treatise on the phenomenon, see LEON FESTINGER, *A THEORY OF COGNITIVE DISSONANCE* (1957). For a brief, and relatively recent, review of the phenomenon, see Richard E. Petty & Duane T. Wegener, *Attitude Change: Multiple Roles for Persuasion Variables*, in THE HANDBOOK OF SOCIAL PSYCHOLOGY, 4th ed. 282, 335-337 (D.T. Gilbert, S.T. Fiske & G. Linzey, eds., 1998). Where one’s beliefs are inconsistent with one’s actions, one can eliminate the inconsistency by changing either the belief or the behavior.

⁸³ Gordon Hodson, Gregory R. Maio, & Victoria M. Esses, *The Role of Attitudinal Ambivalence in Susceptibility to Consensus Information*, 23 BASIC & APP. SOC. PSYCH. 197 (2001).

quality) decisions. Unfortunately, research does not support this belief.

First, ambivalent attitudes don't predict behaviors well at all.⁸⁴ The moral inclusionary model, by inducing ambivalence in judges, introduces "noise" into decisions. This means, for instance, that like cases are less likely to be treated alike across trials: in one hearing, being a child soldier might lead to a judge deciding that an applicant for asylum could receive it after all; in another, even with an identical story and even an identical judge, it might not.

Second, while ambivalent attitudes introduce variance, they also introduce some bias. As already discussed, inducing empathy in an individual predictably inspires an emotional response of sympathy and compassion.⁸⁵ Systematically, sympathy toward a target will tend to produce behaviors favoring that target. Again, advocates of the moral inclusionary model might not object: where the story of the person being judged *should* mitigate the desire to punish him, the model is doing a good job. The problem is that the model is a blunt instrument. It will not, of course, always be the case that telling the full narrative (like Goethe's hypothetical sad childhood, or Amadu's own painfully real history) will result in the judge exonerating them, but telling it does put a thumb on the scale. At the margin, we would expect not only that some who would otherwise be found guilty would go free, but also that some guilty would not be adjudged as culpable—a softening of attitudes toward the wrongdoer that would also have real consequences.

Research in psychology can't, of course, prove (or even disprove) the point that inducing empathy is good, since such a claim is distinctly normative. However, research does describe what happens when such empathy occurs, and this research describes yet another practical problem that the moral inclusionary model must face. In this next section, I describe how studies on "perspective taking" illustrate the promise of empathy towards others, but also hint at the potential perils of empathy should the target "Other" be someone who deserves contempt.

From the outset, note that in order to induce a judge to take the perspective of another (such as a criminal defendant), the judge must be instructed to imagine the other's point of view. In the setting of a psychology experiment, this is generally achieved straightforwardly, for

⁸⁴ Mark Conner, Rachel Povey, Paul Sparks, Rhiannon James & Richard Sheperd, *Moderating Role of Attitudinal Ambivalence Within the Theory of Planned Behaviour*, 42 BRIT. J. SOC. PSYCH. 75 (1996).

⁸⁵ C. Daniel Batson, Judy G. Batson, Carl A. Griffitt, Sergio Barrientos, J. Randall Brandt, Peter Sprengelmeyer, & Michael J. Bayly, *Negative-State Relief and the Empathy-Altruism Hypothesis*, 56 J. PERS. & SOC. PSYCH. 922 (1989); Mark H. Davis, *Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach*, 44 J. PERS. & SOC. PSYCH. 113 (1983)..

example, the experimenter might instruct the participant to “imagine how the person ... feels about what has happened and how it has affected his or her life,”⁸⁶ or perhaps “imagine a day in the life of this individual as if you were that person, looking at the world through his eyes and walking through the world in his shoes.”⁸⁷ In a novel, the perspective-taking “manipulation” happens organically, as the character’s feelings, experiences and circumstances are richly detailed and the reader is immersed in them. In a criminal trial, perspective-taking can be induced in either way: a lawyer (or judge) could explicitly urge the judge to consider the perspective of the defendant before rendering a decision, or perspective-taking could be induced more organically, through offering rich evidence out of which the judge can draw narrative sense.⁸⁸

Urging people to take the perspective of another in the normal run of things yields many positive consequences. Emotionally, it induces compassion and sympathy in the perspective-taker. This has profoundly beneficial effects for both the giver of the sympathy; the brains of the world’s best-trained compassionate thinkers—Buddhist monks, who spend tens of thousands of hours throughout their lifetime engaging in “compassionate meditation”—strongly exhibit brain patterns linked with happiness during functional magnetic resonance imaging studies.⁸⁹ The tendency to take the perspective of others has also been positively associated with self-esteem and social competence.⁹⁰

But it benefits the receiver of sympathy as well. Empathy towards another induces acts of altruism, leading to a reallocation of resources (in experiments, these have included raffle tickets⁹¹, better jobs,⁹² volunteering

⁸⁶ C. Daniel Batson, Shannon Early & Geiovanni Salvarani, *Perspective Taking: Imagining How Another Feels Versus Imagining How You Would Feel*, 751 PERS. & SOC. PSYCH. BULL. 753 (1997).

⁸⁷ Adam D. Galinsky & Gordon B. Moskowitz, *Perspective-Taking: Decreasing Stereotype Expression, Stereotype Accessibility, and In-Group Favoritism*, 78 J. PERS. & SOC. PSYCH. 708, 711 (2000).

⁸⁸ Cf. Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 521 (1991); Nancy Pennington & Reid Hastie, *Evidence Evaluation in Complex Decision Making*, 51 J. PERSONALITY & SOC. PSYCHOL. 242 (1986).

⁸⁹ Antoine Lutz, Julie Brefczynski-Lewis, Tom Johnstone, & Richard J. Davidson, *Regulation of the Neural Circuitry of Emotion by Compassion Meditation Effects of Meditative Expertise*, PLoS ONE 3(3): e1897. Available online at <http://www.plosone.org/article/lookup?articleURI=info:doi/10.1371/journal.pone.0001897> (last visited xxxx).

⁹⁰ Mark H. Davis, *Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach*, 44 J. PERS. & SOC. PSYCH. 113 (1983).

⁹¹ C. Daniel Batson, Judy G. Batson, R. Matthew Todd, Beverly H. Brummett, Laura L. Shaw & Carlo M. R. Aldeguer, *Empathy and the Collective Good: Caring for One of the Others in a Social Dilemma*, 68 J. PERS. & SOC. PSYCH. 619, 626 (1995).

time,⁹³ and the like) towards the source of empathy.⁹⁴ Empathizers are also less likely to respond aggressively against the other,⁹⁵ even in the face of attack.⁹⁶ When describing the object of empathy, moreover, perspective-taking causes perceivers to engage in less stereotyping, an effect that persists beyond the initial task where the perceiver was instructed to take the target's perspective.⁹⁷ Finally, perspective-takers are less likely to suffer from "correspondence bias,"⁹⁸ in which we assume that others' behaviors are a product of immutable personality traits rather than of their circumstances, even though we give ourselves the benefit of the doubt for the same behaviors, believing our own actions are governed by the situations we find ourselves in rather than because we are "that kind of person." That is, perspective takers are more likely to offer explanations for the other's behavior, that they would have offered for themselves.⁹⁹

That is the upside of perspective-taking and empathy. But research has also hinted at a disturbing downside. For one thing, empathizers will sometimes behave altruistically towards a target not only when doing so interferes with their own self-interest, but even when doing so reduces the

⁹² C. Daniel Batson, Tricia R. Klein, Lori Highberger & Laura L. Shaw, *Immorality from Empathy-Induced Altruism: When Compassion and Justice Conflict*, 68 J. PERS. & SOC. PSYCH. 1042, 1045 (1995).

⁹³ Patricia A. Oswald, *Does the Interpersonal Reactivity Index Perspective-Taking Scale Predict Who Will Volunteer Time to Counsel Adults Entering College?*, 97 PERCEPTUAL & MOTOR SKILLS 1184 (1996).

⁹⁴ See generally C. Daniel Batson, *Altruism and Prosocial Behavior*, in THE HANDBOOK OF SOCIAL PSYCHOLOGY 282, 4th ed. (D.T. Gilbert, S.T. Fiske & G. Linzey, eds. 1998); C. DANIEL BATSON, *THE ALTRUISM QUESTION: TOWARD A SOCIAL PSYCHOLOGICAL ANSWER* (1991); Nancy Eisenberg & Paul A. Miller, *Empathy and Prosocial Behavior*, 101 PSYCH. BULL. 91 (1987).

⁹⁵ Deborah R. Richardson, Georgina S. Hammock, Stephen M. Smith, Wendi Gardner & Manuel Signo, *Empathy as a Cognitive Inhibitor of Interpersonal Aggression*, 20 AGGRESSIVE BEHAV. 275 (1994); Kaj Bjorkqvist, Karin Osterman & Ari Kankiainen, *Social Intelligence – Empathy = Aggression?*, 5 AGGRESSION & VIOLENT BEHAV. 429 (2000).

⁹⁶ Deborah R. Richardson, Laura R. Green & Tania Lago, *The Relationship Between Perspective-Taking and Nonaggressive Responding in the Face of an Attack*, 66 J. PERSONALITY 235 (1998).

⁹⁷ See generally Galinsky & Moskowitz, *Perspective-Taking*. In contrast, admonishing the perceiver not to stereotype results in rebound effects, where post-task, the perceiver stereotypes the target *more* than if given no instructions at all. *Id.* See also Adam D. Galinsky & Gordon B. Moskowitz, *Further Ironies of Suppression: Stereotype and Counterstereotype Accessibility Following Suppression*, 43 J. EXP. SOC. PSYCH. 833 (2007).

⁹⁸ Daniel T. Gilbert, *Ordinary Personology*, in THE HANDBOOK OF SOCIAL PSYCHOLOGY 89, 4th ed. (D.T. Gilbert, S.T. Fiske & G. Linzey, eds. 1998).

⁹⁹ R.E. Galper, *Turning Observers into Actors: Differential Causal Attributions as a Function of 'Empathy,'* 10 J. RESEARCH IN PERSONALITY 328 (1976); D.T. Regan & J. Totten, *Empathy & Attribution: Turning Observers Into Actors*, 32 J. PERS. & SOC. PSYCH. 860 (1975).

collective good of other group members.¹⁰⁰ Indeed, empathizers will do this even when they self-consciously acknowledge that their partiality violates principles of fairness and justice.¹⁰¹ In this context, the moral inclusionary model would urge benevolence (toward the offender) at the expense of justice (toward the victim and the community).

C. Daniel Batson, perhaps the most prominent researcher in the psychology of altruism, and his colleagues have detailed the four conditions under which individuals induced to empathize with an other *might* behave altruistically towards him at the expense of their own moral principles:

(a) The person is required to perform behavior for which a personally held moral principle, such as a principle of justice, is clearly relevant; (b) empathy is felt for one or more of the individuals affected by the behavior; (c) the individual or individuals for whom empathy is felt are not the only ones in need; and (d) not everyone's need can be met. In such a situation, empathy-induced altruism should motivate the person to show partiality toward the individual(s) for whom empathy is felt, even if doing so violates the person's moral principles. If the altruistic motive is stronger than the moral motive, the person should act immorally.¹⁰²

These conditions are present in the moral inclusionary model of passing judgment on a criminal wrongdoer. First, the judge is asked to apply a principle of justice—rendering a verdict on the basis of actual guilt or innocence. Second, the judge is invited, by hearing the full story of the criminal defendant in all its narrative richness, to feel empathy towards the one on whom judgment must be passed. Third, the criminal defendant is not the only individual “in need”—indeed, the whole purpose of a criminal trial is to gain vindication for victims, and peace, resolution and order for a community. And fourth, one cannot pass judgment on the wrongdoer without also making a statement about the value of the victim's and community's claims against the defendant. It is generally a zero-sum game—either the victim and the community are vindicated, or the defendant is. Of course, Batson *et al.* do not say that feeling empathy will *always* result in altruistic behavior towards the one at the expense of justice towards the other—this will only happen “[i]f the altruistic motive is stronger than the moral motive.” But this qualification should not make us sanguine—for one thing, it implies that in close cases, a judge *would* tip the balance in favor of the defendant at the expense of justice.

¹⁰⁰ Batson, Batson, Todd, Brummett, Shaw & Aldeguer, *Empathy and the Collective Good* at 626, 628-629. In these experiments, the sympathetic “other” was a fellow group member.

¹⁰¹ Batson, Klein, Highberger & Shaw, *Immorality from Empathy-Induced Altruism* at 1046-47, 1050-51.

¹⁰² Batson, Klein, Highberger & Shaw *Immorality from Empathy-Induced Altruism* at 1043.

If this was the worst possible consequence of the moral inclusionary model, perhaps it would be a price worth paying. Unfortunately, occasional injustice towards victims and communities in favor of altruism towards offenders is not the only consequence the moral inclusionary model must contend with. Perspective-taking has cognitive consequences for those who engage in it, as well. Why does perspective taking lead to altruism, less aggression, greater benefit-of-the-doubt giving, and reduced reliance on stereotyping? Because by putting oneself in another's shoes, in a very real psychological sense, one *becomes* that other person—our cognitive representation of our self gets merged with our representation of the other.¹⁰³ Simply put, perspective-taking makes us perceive the other as more similar to ourselves.¹⁰⁴

A long line of research on juries (and sometimes judges) suggests the risk inherent here: juries are less likely to convict, or are more reluctant to punish harshly, defendants they perceive as similar to themselves. The similarity effect has been studied for sharing the same race,¹⁰⁵ gender¹⁰⁶ and socioeconomic status.¹⁰⁷ It is important to note that these studies don't show simple racism towards minorities, sexism towards women (or men), or bias against poor defendants. The studies show *interactive* effects: women are less likely to convict or harshly punish women; blacks are less likely to convict or harshly punish blacks; the poor are less likely to convict or harshly punish the poor. Thus it would be hard to advocate the moral

¹⁰³ See generally Mark H. Davis, Laura Conklin, Amy Smith & Carol Luce, *Effect of Perspective Taking on the Cognitive Representation of Persons: A Merging of Self and Other*, 70 J. PERS. & SOC. PSYCH. 713 (1996); Galinsky & Moskowitz, *Perspective-Taking* at 715-716. But see C. Daniel Batson, Karen Sager, Eric Garst, Misook Kang, Kostia Rubchinsky & Karen Dawson, *Is Empathy-Induced Helping Due to Self-Other Merging?*, 73 J. PERS. & SOC. PSYCH. 495 (1997) (arguing that true altruism, that is not egoistically-motivated to reduce the giver's own distress, requires cognitive separation with the object of empathy).

¹⁰⁴ Galinsky & Moskowitz, *Perspective-Taking*; Davis, Conklin, Smith & Luce, *Effect of Perspective Taking on the Cognitive Representation of Persons*; John A. Bargh, Mark Chen & Lara Burrows, *Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Activation on Action*, 71 J. PERS. & SOC. PSYCH. 230 (1996).

¹⁰⁵ Howard C. Daudisiel, Harmon M. Hosch, Malcolm D. Holmes & Joseph B. Graves, *Effects of Defendant Ethnicity on Juries' Dispositions of Felony Cases*, 29 J. APPLIED SOC. PSYCH. 317 (1999); Ramsey McGowen & Glen D. King, *Effects of Authoritarian, Anti-Authoritarian, and Egalitarian Legal Attitudes on Mock Juror and Jury Decisions*, 51 PSYCH. REPORTS 1067 (1982); Derek Chadee, *Race, Trial Evidence and Jury Decision Making*, 1 CARIBBEAN J. CRIMINOLOGY & SOC. PSYCH 59 (1996).

¹⁰⁶ Stuart Nagel & Lenore J. Weitzman, *Double Standard of American Justice*, 9 SOCIETY 171 (1972); Gloria J. Fischer, *Gender Effects on Individual Verdicts and on Mock Jury Verdicts in a Simulated Acquaintance Rape Trial*, 36 SEX ROLES 491 (1997).

¹⁰⁷ Freda Adler, *Socioeconomic Factors Influencing Jury Verdicts*, 3 N.Y.U. REV. L. & SOC. CHANGE 1 (1973).

inclusionary model as a simple corrective to a racist, sexist, classist criminal justice system: where judges already are biased in favor of members of their ingroups, the moral inclusionary model would make this bias even *worse*. Another perverse feature of the similarity bias is that it occurs is that it really only occurs when the evidence against the defendant is weak; when the evidence against the defendant is strong, similarity with the defendant leads jurors to be *more* harsh toward him.¹⁰⁸ In other words, the moral inclusionary model is most likely to tip the stack in favor of the defendant in close cases—which is exactly when we would want judges to be least affected by any sort of bias. (When evidence against the defendant is very strong, the similarity effects brought about by the moral inclusionary model simply introduce a piling-on effect against the defendant—this is known as the “black sheep effect” in psychology.¹⁰⁹)

The mechanism by which perspective-taking leads to the perception of similarity is worth discussing in its own right: perspective-takers tend to describe themselves and the target as sharing personality traits; that is, they “confuse” the representation of “self” with the representation of “other.”¹¹⁰ This confusion could occur in two directions: features of the self could be misattributed to the other—that is, if I see myself as funny, intelligent, helpful, disorganized and fearful, after perspective-taking I ascribe those characteristics to the other more than I would have without perspective-taking.¹¹¹ But the blending could occur in the other direction: before I engaged in perspective taking, I might have thought that the other was funny, intelligent, helpful, disorganized and fearful, but after perspective-taking I ascribe those traits to myself more.¹¹²

¹⁰⁸ Chadee, *Race, Trial Evidence and Jury Decision Making*; Norbert L. Kerr et al., *Defendant-Juror Similarity and Mock Juror Judgments*, 19 L. & HUM. BEHAV. 545; Nancy J. King, *Postconviction Review of Jury Discrimination: Measuring the Effects of Juror Race on Jury Decisions*, 92 MICH. L. REV. 63, 82-85 (1993).

¹⁰⁹ Jose M. Marques, Vincent Y. Yzerbyt * Jacques-Phillippe Leyens, *The ‘Black Sheep Effect’: Extremity of Judgments Towards Ingroup Members as a Function of Group Identification*, 18 EUROPEAN J. SOC. PSYCH. 1 (1988).

¹¹⁰ Galinsky & Moskowitz, *Perspective-Taking*; Davis, Conklin, Smith & Luce, *Effect of Perspective Taking on the Cognitive Representation of Persons*.

¹¹¹ Galsinky & Moskowitz, *Perspective-Taking*; Davis, Conklin, Smith & Luce, *Effect of Perspective Taking on the Cognitive Representation of Persons*.

¹¹² Arthur Aron, Elain N. Aron, Michael Tudor & Greg Nelson, *Close Relationships as Including Other in the Self*, 60 J. PERS. & SOC. PSYCH. 241 (1991); cf. John A. Bargh, Mark Chen & Lara Burrows, *Automaticity of Social Behavior: Direct Effects of Train Construct and Stereotype Activation on Action*, 71 J. PERS. & SOC. PSYCH. 230 (1996).

The conditions under which process will dominate are unclear,¹¹³ but no matter which process is in play, they are potentially disturbing in the context of the criminal law. If judges perceive the targets to have similar—generally positive¹¹⁴—personality traits to their own, then *if* defendants on the whole *do not* share the features of the average person, then defendants are being perceived inaccurately. Perhaps one of the points of the moral inclusionary model is that people—even members of despised outgroups—are more alike than different. When the outgroup in question are African-Americans, or the elderly, or even some outgroup about which we know almost nothing,¹¹⁵ this seems like a safe and benevolent assumption. It seems less so when we are talking about criminal defendants who, if guilty, have behaved in often deeply anti-social ways.

This point is even sharper to the extent the judge misattributes the target's features as his own. If the target is a lawbreaker, then the consequence might be that it becomes easier for judges to imagine themselves behaving in anti-social ways. Dan Kahan has written about this phenomenon by arguing that a primary reason individuals obey the law is that they perceive that others like them obey, too.¹¹⁶ When they believe that similar others *do not* follow the rules, they become less likely to do so themselves in order to avoid feeling like “chumps.”¹¹⁷ That is, to the extent a criminal lawbreaker seems less like a distinct, dissimilar “other,” the more her behavior could will appear to be normal, or even sensible. Consider the findings of the Minnesota Tax Experiment—the largest predictor of tax compliance versus cheating was not perceived penalties or likelihood of getting caught, but whether or not the taxpayer thought that other citizens in his or her bracket were taking unjustified deductions.¹¹⁸

There is one more consequences of perspective-taking worth noting: judges actually like the target more than if they had not engaged in

¹¹³ Though guesses have been made; namely, the longer one knows the “other” and the more integrated into one's life, the more merging should result. Aron, Aron, Tudor & Nelson, *Close Relationships as Including Other in the Self*; Davis, Conklin, Smith & Luce, *Effect of Perspective Taking on the Cognitive Representation of Persons*.

¹¹⁴ Davis, Conklin, Smith & Luce, *Effect of Perspective Taking on the Cognitive Representation of Persons*.

¹¹⁵ Galinsky & Moscovitz, *Perspective-Taking*.

¹¹⁶ Dan M. Kahan, *The Logic of Reciprocity: Trust, Collective Action, and the Law*, 102 Mich. L. Rev. 71 (2003).

¹¹⁷ Kahan, *Reciprocity* at 85.

¹¹⁸ Marsha Blumenthal, Charlie Christian & Joel Slemrod, *Do Normative Appeals Affect Tax Compliance? Evidence From a Controlled Experiment in Minnesota*, 54 NAT'L TAX J. 125 (2001); Joel Slemrod, Marsha Blumenthal & Charles Christian, *Taxpayer Response to an Increased Probability of Audit: Evidence From a Controlled Experiment in Minnesota*, 79 J. PUB. ECON. 455 (2001).

perspective-taking.¹¹⁹ Again, proponents of the moral inclusionary model might argue that this is not an unfortunate side-effect, but one of the points in its favor. And if liking the defendant more does not decrease the probability that judges will appropriately condemn a wrongdoer, at least from a behavioral point of view there is little reason to criticize the moral inclusionary model on this ground. However, from an expressive point of view, making judges more fond of the object of their condemnation seems problematic. Even engaging in the reciprocal relationship that exists between “author” (defendant) and “reader” (judge) invites a kind of friendship—albeit a temporary one, in much the same way reading a book fashions a brief friendship with the characters in the novel—which might rightly make us deeply uncomfortable.¹²⁰

The arguments I have presented so far do not say that the moral inclusionary model is *wrong*. Instead, I have made two substantial modifications to it: the first is to lay out more precisely what the model demands of judges, and the benefits it actually promises. The second is to offer skepticism that the positive cognitive and emotional effects of moral inclusion will happen consistently and in the way proponents would like, and to warn of the possible perverse effects that will often accompany the application of the model. I turn now to applying my contributions to the moral inclusionary model, to the law itself.

Part II. Moral Exclusion in Law: Applying the Lessons

Those who would have us apply narrative thinking to the task of judging should not lose the exclusionary lessons of literature and psychology. In order to do a good job of judging, the judge must be cautioned not just to include morally relevant particulars (a familiar and fairly uncontroversial theme), but purposely and carefully to *exclude* morally *irrelevant* particulars. As in literature, good judging is a process of filtration. Some information deserves a judge’s attentive consideration, but some may either be irrelevant to a just resolution of a case, or in fact *preclude* a just resolution.

The law currently embraces the moral exclusionary view, requiring that judges actively enable filtering mechanisms in their decision making. I have already mentioned the most vivid example of this: the role judges have in barring evidence from the view of jurors at trial. Consider, for example, the ban on “past bad acts” evidence offered to show that a party has a character-

¹¹⁹ Galinsky & Moscovitz, *Perspective-Taking: Davis, Conklin, Smith & Luce, Effect of Perspective Taking on the Cognitive Representation of Persons*.

¹²⁰ The idea of reading as forming a sort of friendship with the implied author comes from Wayne Booth. But as already mentioned, Booth focuses on why we should invite such friendships, with some minor qualifications suggesting there might be times we reject them—but seemingly not without giving them a try first. This is consistent with the moral inclusionary model, in whose camp Booth is solidly at home. BOOTH, *THE COMPANY WE KEEP* at 201-223.

based propensity to engage in bad behavior.¹²¹ The idea that the bar on character-propensity evidence advances the accuracy of the ultimate verdict is fairly laughable.¹²² If a defendant has robbed a bank fifteen times before, then it simply is more likely that he robbed the bank this time rather than being the unfortunate victim of mistaken identity;¹²³ if a defendant's last two wives were found drowned in a bathtub, it is more probable that his third wife found drowned in the tub was murdered rather than having drowned by accident.¹²⁴ Moreover, if the defendant had only robbed a bank once before, say long ago when he was a foolish youth of eighteen, or if the previous wives had drowned not in a tub but while swimming at a public beach in one case and in a pool after suffering an aneurism in the other, then a judge could always exclude the evidence for being insufficiently probative—that is, judges could always exclude weak or prejudicial past-acts evidence on a case-by-case basis.

But that is not the rule we have. The character-propensity bar is categorical, meaning the evidence gets excluded whether or not it is accuracy-enhancing in a given case. Why? The rule can only plausibly be classified as “morally”-based, meaning not that exclusion improves verdicts in some probabilistic sense, but improves them in some legitimacy-enhancing or expressive sense. Various morality- (as opposed to accuracy-) based justifications for the rule have been proposed, most of which reduce to the idea that the rule expresses a commitment to respecting individual free-choice and autonomy *not* to behave in ways a defendant might have behaved in the past.¹²⁵ But the character-propensity rule is not unusual in this sense. Many—perhaps most—of the categorical exclusion rules function to keep legitimacy-reducing, rather than simply accuracy-reducing, evidence away from jurors.

Of course, you may not agree that exclusion of character-propensity evidence *does* enhance the legitimacy of verdicts. After all, the autonomy of the accused is purchased at the price of a reduced likelihood of vindication for the victim. Indeed, ambivalence about the value of expressing commitment to autonomy for defendants surely explains why the character-propensity rule is both riddled with explicit exceptions,¹²⁶ and is so widely

¹²¹ This is an old common-law rule, and is codified in the federal system as FEDERAL RULE OF EVIDENCE 404.

¹²² See generally Susan Marlene Davies, *Evidence of Character to Prove Conduct: A Reassessment of Relevancy*, 27 CRIM. L. BULL. 504 (1991).

¹²³ *United States v. Danzey*, 594 F.2d 905 (2d Cir. 1979)

¹²⁴ *Rex v. Smith*, 11 Cr. App. R. 229, 84 L. J. K. B. 2153 (1915).

¹²⁵ Gerald E. Lynch, *RICO: The Crime of Being a Criminal, Parts III & IV*, 87 COLUM. L. REV. 920, 936 (1987).

¹²⁶ See, e.g. FEDERAL RULE OF EVIDENCE 406 (exempting from categorical exclusion behaviors and actions that can be described as “habits”); the exceptions listed inside FEDERAL RULE OF EVIDENCE 404 itself, under section 404(a)(1)-(3); and the exceptions to the exceptions described by FEDERAL RULES OF EVIDENCE 412-415.

evaded in practice via paper-thin arguments in particular cases (such as that such evidence is being offered not to show “propensity” but for some “other purpose.”)¹²⁷ But even if one agrees that the character-propensity bar does in fact advance the good, that does nothing to take away from the notion that whether one wishes to include or to bar requires a *moral* evaluation.

As Larry Tribe has written,

It would be a terrible mistake to forget that a typical lawsuit, whether civil or criminal, is only in part an objective search for historical truth. It is also, and no less importantly, a ritual — a complex pattern of gestures comprising what Henry Hart and John McNaughton once called “society’s last line of defense in the indispensable effort to secure the peaceful settlement of social conflicts.” One element, at least, of that ritual of conflict-settlement is the presence and functioning of the jury — a cumbersome and imperfect institution, to be sure, but an institution well calculated, at least potentially, to mediate between “the law” in the abstract and the human needs of those affected by it.¹²⁸

In a similar vein, Charles Nesson argued,

The goal of generating acceptable verdicts is not met simply by choosing the verdict that is most probably accurate. Acceptable verdicts and probable verdicts might appear to coincide, given that one obvious way to gain public acceptance is to search for truth. But the correlation between probability and acceptability is not exact: a probable verdict may not be acceptable, and an acceptable verdict may not be probable.¹²⁹

In this section I wish to further explore this observation about the acceptability of moral exclusion as it is already practiced in law right now. I start with an example drawn from case law that serves as a real-life parallel to the moral exclusion practiced by Thomas Keneally in *Schindler’s List*. I then address objections to this example, which focus on the idea that the world of literature is and should be different from the world of law when it comes to moral exclusion. Finally, I discuss what is perhaps the deepest challenge to my endorsement of the moral exclusionary model of the law: the idea that moral exclusion view does and should apply in *all* areas of the law, including capital sentencing.

A. An example from case law

¹²⁷ See the list of “other purposes” for which past acts can be shown, in FEDERAL RULE OF EVIDENCE 404(b). David J. Karp, *Evidence of Propensity and Probability in Sex Offense Cases and Other Cases*, 70 CHI.-KENT L. REV. 15 (1994).

¹²⁸ Lawrence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971).

¹²⁹ Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985).

Dan M. Kahan and Martha Nussbaum, in their article “Two Conceptions of Emotion in Criminal Law,” relate the case of *Commonwealth v. Carr*.¹³⁰ There, a man, accused of killing two lesbians he observed engaging in sex, was denied the right to introduce psychiatric evidence that would have demonstrated that he had experienced an uncontrollable urge to kill stemming from his deep, clinical homophobia.¹³¹ Carr’s story is the legal parallel to Goethe’s. If the law and literature movement as it is usually described were taken at face value, then just as with Goethe, it would *seem* as though these scholars would demand that Carr’s story be told. Again, he is as human as you and me; and recognizing him as such demands that this “hard case” be heard.

But this can’t be right. Just as we couldn’t fully empathize with both Goethe and the Schindlerjuden, our sympathies for both Carr and his two victims are simply incompatible. And just as with Goethe in *Schindler’s List*, we cannot hide behind the notion that Carr’s story is “structurally” unnecessary. Kahan and Nussbaum argue convincingly that the basis for this exclusion was not that the evidence was irrelevant to determining whether or not Carr was genuinely provoked to kill, as it assuredly was. Carr wanted the jury to have the option of finding him guilty of the lesser charge of manslaughter rather than murder; the difference is that the manslaughter defense requires proof that the defendant, “at the time of the killing... is acting under a sudden and intense passion resulting from serious provocation.”¹³² Instead, the exclusion reflected an assessment that Carr’s reaction was morally outrageous, because “[a] reasonable person would simply have discontinued his observation and left the scene”.¹³³

But as Kahan and Nussbaum point out, at the same time we reject Carr’s inclusion of his psychiatric evidence for moral reasons, we *might* include another defendant’s psychiatric evidence of provocation in a different situation, say, that of a man who kills his wife after learning she is pregnant with his stepbrother’s child,¹³⁴ or a woman who kills her batterer after years of abuse.¹³⁵ The difference is not in the degree of reliability of the evidence in demonstrating a mentally disabling condition, evidence that the court, were it value neutral, would have to agree was relevant as a matter of law. The distinction instead focuses on the degree to which the law thinks that a full accounting of the defendant’s point of view is morally justified.

¹³⁰ 580 A.2d 1362 (Pa. Super. Ct. 1990)

¹³¹ Dan M. Kahan and Martha C. Nussbaum, *Two Conceptions of Emotion in Criminal Law*, 96 COLUM. L. REV. 269, 272 (1996).

¹³² 18 Pa. C. S. §2503(a).

¹³³ 580 A.2d at 1364.

¹³⁴ *Commonwealth v. McCusker*, 448 Pa. 382, 292 A.2d 286 (1972).

¹³⁵ *Commonwealth v. Stonehouse*, 521 Pa. 41, 555 A.2d 772 (1989).

B. Objections to the moral exclusionary model in law

Of course, just because this is the way the law is currently does not mean this is how it should be. That such filtering makes for a better-reading and more ethically sound literature does not mean that it makes for better criminal law. One oughtn't take the comparison of law to literature too far, this argument goes, since the consequences of encountering a bad book are pretty limited (you've maybe just wasted a little time), while the consequences of bad criminal judging are dangerous in the extreme. Just because good literature can't exist without a fair balance between moral inclusion and moral exclusion, doesn't mean that good *law* ever demands that moral inclusion be tempered by moral exclusion of the sort the judge in *Carr* engaged in.

There are two types of critics who might object to applying the moral exclusionary model of literature to the law. The first I will dismiss rather summarily. He is the anti-moralist critic, who rejects the notion that the law (unlike, perhaps, literature) should pay any attention at all to "moral" or expressive considerations. Judge Richard Posner is the most clear-cut critic in this camp;¹³⁶ but his moral agnosticism was shared by perhaps the most prominent theorists of twentieth-century American legal thought, Justice Oliver Wendell Holmes:

For my own part, I often doubt whether it would not be a gain if every word of moral significance could be banished from the law altogether, and other words adopted which should convey the legal ideas uncolored by anything outside the law. We should lose the fossil records of a good deal of history and the majesty got from ethical associations, but by ridding ourselves of an unnecessary confusion we should gain very much in the clearness of our thought.¹³⁷

Maybe this is an admirable aspiration for the law (although I can't say I would agree), but it has been fairly thoroughly discredited in actual practice. One can describe the entire Legal Realist movement as an answer to the viability of Holmesian moral neutrality.¹³⁸ It is hard even to conceive of law absent a sense of morality—the very things we decide to make illegal reflect a shared (or aspired to) vision of ethical behavior and valuations of people

¹³⁶ See generally RICHARD A. POSNER, *LAW AND LITERATURE* (1995).

¹³⁷ Oliver Wendell Holmes, Jr., *The Path of the Law*, 10 HARV. L. REV. 457 (1897).

¹³⁸ See, for example, Robert Hale, *Coercion and Distribution in a Supposedly Noncoercive State*, 38 POL. SCI. Q. 470 (1923); Karl Llewellyn, *Some Realism About Realism—Responding to Dean Pound*, 44 HARV. L. REV. 1222 (1931); Felix Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809 (1935).

and things.¹³⁹ At any rate, I can do little more to address this critic here, that others have not done better elsewhere.

The second type of critic, though, is square within my line of fire. This critic agrees that the law does and must pay attention to moral considerations, but she rejects the propriety of any legal actor categorically, and especially *ex ante*, judging any particular point of view as valueless. She holds a strict moral inclusionary view in *law*, whether or not she does in literature or whatever the lessons of psychology, because of concerns distinctive to the legal context. That is, literature and even psychology have different goals than the law. Literature is aspirational, evocative and provoking—its very essence is choosing sides, and it stands ready to be judged for the choices it makes. As for psychology, perhaps it is very, very hard for individuals to simultaneously and fully condemn and comprehend, perhaps taking the view of a criminal offender has effects that can be potentially corrupting to the perspective-taker and to other goals of the criminal justice system—but this does not end the discussion. Science cannot tell the criminal justice what it ought to do, it can only tell it what the potential costs are.

The law, by contrast, must judge. Guilty verdicts don't just have expressive consequences, or potential psychological consequences, they have *actual* consequences for actual human beings. The law is not literature and the law is not psychology; while these other disciplines can offer insight into what the law does and how it does it, they cannot ultimately answer the normative question of what the law should do. This type of critic has essentially three concerns about moral exclusion, which I will call her fundamental, institutional, and structural doubts. I address each in turn.

1. *Fundamental doubts.* The first objection takes up the perspective of dual compassion. It holds that we are obliged to sympathize with all victims, even the ones who as a result of their past misfortune become victimizers themselves. In other words, if Carr's homophobia is a product of his own victimization (he claims it was a result of "a history of constant rejection by women, including his mother who may have been involved in a lesbian relationship, sexual abuse while in prison in Florida, inability to hold a job, and retreat to the mountains to avoid further rejection"),¹⁴⁰ then we are duty bound to have dual compassion for both him and his two victims, and accordingly, any limitations on his ability to tell the stories which might trigger that compassion would be illegitimate.

I have already reviewed the evidence about the psychological difficulties of dual compassion. Often, sympathies are just mutually incompatible. That is, one cannot feel sympathy for both a wrongdoer and his victim at the same time and in reference to the same crime and still maintain a sharp sense of

¹³⁹ For a complete answer to this critic, see Kahan & Nussbaum, *Two Conceptions of Emotion* at 359-362.

¹⁴⁰ Carr at 1363-1364.

the moral divide between the two. Does this mean, though, that the law must concede the point? The answer must turn on whether or not the law can legitimately demand the psychologically impossible, or even the psychologically difficult. Certainly the law requires that citizens do, or refrain from doing, many things that are psychologically very tough: in no state, for instance, is one allowed to use deadly force to answer even the deepest insult. Just because juries sometimes let off defendants who attack in the face of such offense¹⁴¹ does not undermine the point—the law formally condemns and punishes it, and it would be wrong for the law to allow physical “self defense” to insults, even if *everyone* in a particular jurisdiction would endorse it.

The anti-exclusion argument is grounded in an appeal to the moral legitimacy of the law—to the extent the law indulges the base desire to exclude the full perspective of wrongdoers and treat them as human beings with stories to tell, the law loses its moral authority. This in turn should lead to disrespect for the law’s proclamations, and even its institutions. However, this is not the *only* potential source of illegitimacy in the law. There may be no satisfying response to this argument, because it is essentially exhortation, not argument, and as such admits of no counterproof. But if there is an answer, perhaps it is this: failing to fully condemn when full condemnation is deserved, *also* contributes to ill-regard of the law. In addition, demanding that judges engage in idealistic and unrealistic tasks which they find it difficult or even impossible to do, also undermines the legitimacy of the law. Since the law is damned if it does allow wrongdoers a full narrative voice, and also damned if it doesn’t, this fundamental objection to the moral exclusionary view is at best a philosophical wash.

2. *Institutional doubts.* Even if she accepts the fundamental point about incompatible sympathies, the anti-exclusionary critic might still make an institutional argument and object strongly to *judges* being the ones to choose among those competing sympathies. For instance, in the *Carr* example, Kahan and Nussbaum freely admit there is nothing in their account that dictates that evidence about homophobia will always be excluded, while more morally acceptable evidence (like, depending on your views, expert testimony on battered women’s syndrome) will be included.¹⁴²

Thus, one might argue, contrary to Kahan and Nussbaum, that we should include all important evidence, and leave the moral bottom line to be drawn by whomever has been given the authority to be the ultimate decision maker, which in this example is the jury. If a particular point of view is morally irrelevant, as Carr’s homophobia is, let the final decision maker be the one to decide it is so after hearing the relevant evidence. The judge should not be

¹⁴¹ Consider, for example, the words of a Southern juror who was tried for murder after shooting three other men who had verbally taunted him: “Good God Almighty, bub.... He ain’t guilty. He woudn’t of been much of a man if he hadn’t shot them fellows.” HODDING CARTER, *SOUTHERN LEGACY* 50 (1950).

¹⁴² Kahan & Nussbaum, *Two Conceptions of Emotion* at 362.

allowed to impose his own opinions about what is *morally* relevant and irrelevant, because of the risk that he might make terrible mistakes. More information, this view contends, can only improve the chances that the final decision maker will come to a morally acceptable conclusion, and excluding information on ethical grounds handicaps the ability of that final arbiter to do his, her or its job. The guilty verdict for Carr is condemnation enough for his homophobic motivations; there is no need to exclude them *ex ante*. Indeed, pulling this away from the final judge actually deprives the system of the opportunity to condemn his motivations via the most legitimate, powerful means: for democratic reasons, it is the jury's role to determine morality in a criminal trial; and we must not presume the jurors cannot come up with a morally justified decision on their own. Even if settled law is right and filtering is a good idea (or at any rate unavoidable), the proper final arbiter of moral relevance should not be judges, but the more obvious representative of the norms and sensibilities of the people at large, juries.

One part of this criticism is a point well taken: we should pay attention not just to what stories get told, but where, when and to whom. Some institutions are able to process full narratives in a way that does not lead to some of the perverse expressive consequences I have outlined—a point I take up later. The jury, however, is not one of them—and this is not (just) because juries might make moral mistakes, too.

As an institutional matter, when deciding where and when (rather than whether) narratives ought to be restricted, one consideration is transparency. If we simply left it to the jury, the message of condemnation for Carr's homophobia is deeply muted. Sometimes it will be clear, but more often it will not. Does a conviction mean the jury rejected the idea that his homophobia was mitigating? Or does it simply mean that though mitigating, the other evidence supporting first degree murder overwhelmed it? Because the jury does not have to give reasons for its decisions, the weight they give any piece of evidence remains hidden. Contrast this with a judge, who, when deciding that Carr's homophobia cannot be admitted, does so in a discrete and defined way. The opinion in *Carr* made it clear, because it had to, that the judge was excluding the evidence because he had found it irrelevant; moreover, unlike juries, judges typically explain why they rule as they do. Moreover, if either the judge *or* the jury can make a mistake by excluding what are in fact morally relevant points of view, then it makes sense to assign the task of deciding which is which to the institution capable of having its decision second-guessed, either on appeal or via the court of public opinion—the judge.¹⁴³

¹⁴³ Kahan and Nussbaum give a telling example of exactly this happening. In Texas in 1988, a state judge gave only a slap on the wrist to a murderer, whose victims were two homosexual men who had solicited the defendant for sex, saying “I put prostitutes and gays at about the same level, and I'd be hard put to give somebody life for killing a prostitute.” In response to a public outcry, three things happened, each more remarkable than the last: first, the judge was first censured; second,

Whoever performs the story-filtering function, it must be done with the awareness that this filtering undeniably entails moral evaluation. Moral decisions are political, and so should be made with a maximum of deliberation and under close scrutiny. Therefore a major consideration about who we want to do the filtering is relative visibility. If we can see the filtering happen, we can repair it if it is done wrongly.¹⁴⁴ Legislatures and judges are quite visible in this regard: when they exclude (or include) a story, it is clear they are doing so, and we can repair their mistakes if needed, through, for instance, appellate reversal or even legislative or political intervention. The black box of the American jury, by contrast and by design, is completely opaque. It is therefore the *least* desirable place to vest power to decide whether to include or exclude a wrongdoer's perspective when deciding on a verdict.

Another institutional concern is whether the act of telling will have offensive implications to the other "characters" in the story in one setting, that it might not have in another. A trial is very much like a novel: there are protagonists and antagonists, and which is which is determined by whose story gets the richest telling. In a trial, when the offender tells the narrative of how he came to do what he did, this invites empathy for him. When the victim gets a richer telling, this invites sympathy for *him*. Indeed, the usual criticism of typical criminal trials is not that offenders get too little opportunity to tell their stories, but that *victims* do. (Perhaps then it is no surprise that the two most important Supreme Court decisions endorsing the importance of the narrative approach to criminal trials have resulted in increased opportunities for victims to tell their stories—I speak here of the Court's allowing Victim Impact Statements at capital sentencing hearings,¹⁴⁵ and the already-mentioned right of prosecutors to not allow defendants to "stipulate" away a prosecutor's narrative.¹⁴⁶) In a trial, the mere act of including information about Carr's homophobia would have been a statement in itself, as the judge recognized. Even beyond that, the words used to describe it to the jury, both by counsel and more importantly, by the judge's charge to them, would necessarily include value judgments, as Stanley Fish would be the first to point out.

By contrast, when we allow full, rich narratives in other settings, the tragic trade-off between sympathy for victims and sympathy for the wrongdoer is substantially muted. Which other settings? Legislatures, for one. By allowing those affected by legal judgments to tell their full stories at the wholesale stage, where generally-applicable rules are being fashioned, we get most of the advantages of allowing them to tell their tales at the retail,

citizens voted him out of office; and third, the Texas legislature passed a law increasing the penalties for crimes motivated by hate. See Kahan & Nussbaum, *Two Conceptions of Emotion* at 364.

¹⁴⁴ Kahan & Nussbaum, *Two Conceptions of Emotion* at 362.

¹⁴⁵ *Payne v. Tennessee*, 501 U.S. 808 (1991).

¹⁴⁶ *Old Chief v. United States*, 519 U.S. 172 (1997).

trial level, while avoiding most of the disadvantages. Even if we accept the moral inclusionary model in general, we don't have to accept it at every level of decision making in the law, including in the courtroom.

It is hard to deny the moral force of the model. Not only is the idea of letting those we are judging tell their whole stories an intuitively appealing proposition on its face, but if we were to categorically and always close ourselves off to hearing their stories, we would be inviting moral stagnation. Consider, for instance, the Battered Woman's Syndrome. When defendants first began trying to introduce this evidence, it was overwhelmingly resisted by judges, and systemic denial of the defense to criminal defendants helped to further entrench outdated gender norms about the appropriate balance of power in relationships. But state legislatures had the power to step in and rule that, categorically, the defense would be admissible.

As Kahan and Nussbaum demonstrate, legislatures pass categorical rules of exclusion or inclusion frequently.¹⁴⁷ The legislatures' actions can take the form of modifying the formal evidence rules (like the Federal Rules of Evidence in the federal system, and equivalents in each state), or simply writing substantive statutes in such a way to make particular narratives relevant or not (such as the example I previously offered about whether or not reasonable lack of knowledge about the victim's age would be a defense to statutory rape, or whether or not being conscripted as a child into a militia could negate the usual rule that applicants who persecute will be denied asylum). When they debate such actions, legislatures can and do apply a fairly unadulterated version of the moral inclusionary model before making their decisions. For instance, Federal Rules of Evidence 412 and 413 are fairly recent additions to the Rules, that categorically bar defendants from introducing evidence of a victim's sexual history, and categorically allow victims to introduce evidence of a defendant's propensity to commit sexual aggression. The House and Senate both put on extensive hearings before ratifying the new rules, inviting the testimony not only of experts, but of individuals whose personal stories could help the legislature to form its moral sense about the issues. Other categorical changes have been spearheaded not by Congress directly, but by administrative committees, whose suggestions are vetted using notice-and-comment rulemaking procedures, which include a period during which any and all are invited to express their concerns and, if they like, tell their own stories.

In sum, the moral inclusionary model does not have to insist that narratives always be available at the retail, trial level. Many of its goals of respectful, humane treatment of criminal defendants, and dignity of the legal system, can be achieved instead at the wholesale level, by legislatures and administrative agencies. By housing the moral inclusionary model in the alternate locations, we mostly—though certainly not completely—avoid the tragic-tradeoff inherent in allowing wrongdoers to tell their stories at trial.

¹⁴⁷ Kahan & Nussbaum, *Two Conceptions of Emotion* at 364.

The decision about whether to exclude a narrative must be made *somewhere*—there is just no getting away from that fundamental fact—and demonstrating that this is so is the primary aim of this paper. At some point, some decision maker must determine what evidence is morally relevant to whether or not a harmdoer should be condemned. Such exclusionary decisions are endemic to both literature and law. Authors make the call in literature, and we can judge them, and their work, for how well they do it. In law, there are many possible decision makers, but some are better than others. We can accept the moral inclusionary model, without accepting that it applies with equal force in every legal setting. The next section discusses another way that the scope of the moral inclusionary model might be limited, in order to save it overall.

3. *Structural doubts.* The final argument against the propriety of moral exclusion in law is also the most intriguing. This view tries to save the “dual compassion” perspective by structuring the law in a way that completely separates the message vindicating the victim from the punishment of the offender. After all, there is no reason that *punishment* has to be the exclusive medium to communicate our message of valuing the victim. The state could implement a dual procedure which would both forcefully assert a sense that the victim has genuinely been hurt, but also completely independently recognize limits the wrongdoer might have faced in deciding whether to commit the crime. This would be an attempt to resolve the problem of mutually incompatible sympathies, such as between Goethe and the Schindlerjuden, Bigger and Bessie, and Carr and the lesbians he shot. In “Two Conceptions of Emotion in Criminal Law,” Kahan and Nussbaum suggest that just such a dual system exists in the law, in that we separately try, and then sentence criminal defendants.¹⁴⁸

But this cure demonstrates a misdiagnosis of the disease. The problem is not institutional, but psychological. The human mind does not in the normal course of things separate condemnation from punishment. Going back to literature, it would be as if Keneally had written another, completely separate, novel after *Schindler’s List*, called perhaps *Goethe’s Demons*. We could read the first and sympathize fully with the Jews, then put the book down. We would then read the second, and to the extent we felt compelled by the tale, sympathize with *Goethe*, and then put *it* down. In this way, Kahan and Nussbaum are effectively arguing, we would not have to compromise our compassion for anyone, for Keneally would have conveyed two different, equally valid and uncontradictory messages of value.

Kahan and Nussbaum claim that mercy does not (necessarily or even usually) mitigate blame, but do not give much explanation of why they assume this. The authors merely assert that a light sentence does not give up the condemnation of wrongful behavior, but is instead just an exercise of “mercy” or “forgiveness” which maintains the assessment of wrongdoing

¹⁴⁸ Kahan & Nussbaum, *Two Conceptions of Emotion* at 368.

that was made in trial stage.¹⁴⁹ Citing Jeffrie Murphy, they argue that “[t]o forgive a friend’s inattention... is not to take back the initial *condemnation* of her behavior; on the contrary, the convention of forgiveness presupposes blame, for if the offending conduct were no longer recognized as wrong, there would be nothing to forgive,”¹⁵⁰ and they then apply this observation to the criminal justice system.

While Kahan and Nussbaum are right that we may forgive transgressors privately, they misstep in asserting that this ability for private forgiveness translates neatly into criminal mercy. Kahan and Nussbaum argue that in the first, condemning stage, “we are largely unconcerned with issues of responsibility for character; we are concerned only to appraise the act for what it is – either good or bad, beautiful or ugly.” It is only at the punishment phase that “we are intensely interested in the offending party’s responsibility[.]” But so conceived, are these really two steps? A criminal act isn’t so ugly if it is done by an irresponsible (for whatever morally recognized reason) person — it lacks that element of insult which Murphy was so careful to describe. A judge does not grant *mercy* in mitigating the sentence of an offending party who is not responsible for his acts. Instead the judge is merely meting out appropriate *justice*—those who are not (as) responsible are not (as) guilty.

Once this notion of mercy is disentangled from that of plain justice, Kahan and Nussbaum’s dual scheme becomes as unhelpful as *Goethe’s Demons*. To the extent that leniency in sentencing reflects the requirements of justice, people will naturally have no complaint about the conceptual “separation” of the two. But to the extent that this leniency reflects true “mercy” of the gratuitous, personal sort experienced when one friend forgives another, the public should complain, and complain loudly.

Kahan and Nussbaum recognize the instability of their bifurcated procedure, noting that “nothing prevents assessments made at [the sentencing phase] from colliding with and completely undermining assessments made in the adjudication of guilt.”¹⁵¹ But their use of the word “colliding” is telling. The word connotes a certain randomness and infrequency. Think of traffic: cars on the road for the most part do not “collide” either predictably or often. (If they collided predictably, the problem could be systematically fixed; if they collided frequently, no one

¹⁴⁹ *Id.*

¹⁵⁰ *Id.*

¹⁵¹ Kahan & Nussbaum, *Two Conceptions of Emotion* at 371. Indeed, with the publication of his article *What Do Alternative Sanctions Mean?* 63 U. CHI. L. REV. 591 (1996), Professor Kahan seems to have abandoned the idea that the line between condemnation and punishment can reliably be drawn. That entire article is an elaboration of the notion “that when society deliberately forgoes answering the wrongdoer through punishment, it risks being perceived as endorsing his valuations; hence the complaint that unduly lenient punishment reveals that the victim is worthless in the eyes of the law.” *Id.* at 598.

would drive.) Their use of the term here implies that an act of leniency that undermined the message of condemnation would be a rare event, like a chance automobile accident. On the contrary, I would argue that it is the norm.

“Token” punishments are inevitably seen by the public in one of two ways. Either they are seen as a correct expression that the crime itself was no crime (or not much of one) at all; or, they regard them as evidence of renegade judges or juries imposing their own twisted sense of justice in contravention of acceptable norms. Examples of the former are the routine slaps on the wrist Wyoming gave to speeders in the days of federally-mandated 55 m.p.h. laws, and the minuscule fines New York City judges impose on small-fry possessors of marijuana. As for the latter, Kahan and Nussbaum offer examples themselves. They mention the case of a judge who inspired the outrage of a community by lightly sentencing the murderer of a homosexual male.¹⁵² In another piece, Kahan describes an outpouring of feminist anger after the light sentencing for a gang-rape of a mentally handicapped girl.¹⁵³ Tellingly, Kahan and Nussbaum offer no examples of cases where a community accepted a lenient sentence as nevertheless fully condemning the criminal. Why? Probably because such cases are few and far between, if they exist at all. Occasional promising examples will generally deconstruct into sentiments that the wrongdoer was not even guilty after all, either because the law itself is unjust, or the wrongdoer was not entirely responsible.

The sentiment seems to be fundamental: justice requires a punishment equal to the degree of our condemnation.¹⁵⁴ Contrary to the suggestion of Kahan and Nussbaum, it’s a rare society indeed that does *not* allow “mercy” in sentencing to collide with and undermine any prior statements of official condemnation. We cannot write two novels in each criminal trial; the public will read it as one.

C. Is death different?

Easily the most difficult challenge to my case is the death penalty. My view, obviously, could be taken to imply that we shouldn’t hear from the likes of Goethe or Carr even in a capital case. Of course, this goes directly against legal precedent. The Supreme Court itself, in the famous case of *Lockett v. Ohio*,¹⁵⁵ expressly demanded that a sentencer must not be denied access to any structurally relevant information which might cause him to

¹⁵² Kahan & Nussbaum, *Two Conceptions of Emotion* at 364, 371.

¹⁵³ Kahan, *What Do Alternative Sanctions Mean?* at 598-99, note 28.

¹⁵⁴ Naturally, a society is not *obliged* to match its punishment to its condemnation. Concerns of utility, politics, or even simple decency might dictate leniency. Sometimes justice just has to take a back seat to other concerns—but this does not in any way dilute the practical reality that full justice has been sacrificed in such instances.

¹⁵⁵ 438 U.S. 586 (1978).

mitigate a capital sentence.¹⁵⁶ In *Lockett* and elsewhere, the Supreme Court has held that the death penalty is qualitatively different than the more standard punishments such as prison, and so subject to different procedural requirements before we may impose it.

The question is, should that difference turn a circumstance which would be morally irrelevant in a non-capital case, now relevant, simply because the death penalty might be imposed? If, as I have tried to demonstrate, both the guilt and the sentencing phases of a criminal trial reflect judgments about the responsibility and culpability of the defendant, and we have decided that some circumstances are simply morally irrelevant to those assessments, does the specter of execution alter that? If a defendant were less culpable because of his homophobia, why shouldn't he be entitled to use it in his case-in-chief as a defense to a charge of robbery and murder at a store owned by homosexuals, rather than just at sentencing (that is, of course, if it is even relevant there)? On the other hand, if his homophobia did *not* make him less culpable, why should we allow him to bring it up in sentencing of a capital case? And again if I am right that the sentence reflects the same evaluation, in practical effect, as the conviction itself, why should a defense be barred at the more important phase of the proceeding (that which establishes guilt or innocence), but admitted after the worst damage has already been done to the defendant (he's already been declared guilty)?

These are questions without good answers. There just doesn't seem to be any principled way to treat capital cases differently than other cases in this respect: if the information is morally repugnant in one area, it is equally so in another. And yet, there is the pull that death is different—it's just not clear why. Perhaps the strong intuition we feel (grounded, maybe, in the qualitatively different sanction to be applied) is enough to give the defendant every benefit of the doubt—though again, it's hard to credit this generosity as sincere, given that during the guilt-phase of the trial, the defendant is given no such concessions.

Conclusion

Born out of the law and literature field, the moral inclusionary model of legal judging has garnered considerable attention not only from scholars, but from the courts. In this article, I have clarified exactly what that model entails; namely, that it is a methodological approach to good judging which demands that a criminal defendant must be allowed to tell his full story. The full telling enables—indeed, demands—that judges empathize with those whom they would judge. The claimed advantage is not—or at least not just—that a full telling might mitigate his blameworthiness, but instead that those who are condemned are treated as fully human, and that those who do the condemning are not debased by the experience.

¹⁵⁶ *Id.* at 605.

As promising as the moral inclusionary approach is, I have also argued that it necessarily entails a flip side: the moral *exclusionary* model. According to the exclusionary model, the law not only must include relevant viewpoints, it must categorically and firmly exclude morally *irrelevant* ones. This insight, too, comes from a textured understanding of literature—even though scholars in the field of law and literature have not emphasized, or even openly acknowledged, this component of the narrative approach to judging. Moreover, the exclusionary model finds considerable support in the science of social psychology, which demonstrates (and perhaps even the impossibility) of the purely inclusionary model—namely, since sympathy is a scarce psychological resource, one cannot simultaneously sympathize with and fully condemn a wrongdoer, and one cannot simultaneously sympathize with a wrongdoer and with his victim.

The law, in fact, already joins the moral inclusionary with the moral exclusionary models of judging. Throughout both the law of evidence and, less explicitly, the substantive criminal law, judges must choose *which* stories to hear. A refusal to choose (that is, a decision to include all viewpoints offered) does not leave the criminal law in a morally neutral position; instead, it means that the law has endorsed sympathy for one viewpoint as the expense of sympathy for another—to be legitimate, the law must pick carefully. While the moral inclusionary model is useful for making us aware that some previously-ignored perspectives should be invited into the law; the moral exclusionary model is useful for reminding us why some were left out in the first place.

But how should we choose? There is no neat or uncontroversial way to decide which particulars are to be excluded and which are not. But in any event, one cannot assume that legitimate and acceptable outcomes to cases will emerge as a consequence of rich narrative methodologies for judging. The outcomes will only be moral when the narrative itself is ethically told, which sometimes will require a thin and poor telling. In other words sometimes justice must be left, or even rendered, blind.

Creating (or even acknowledging) an independently-theorized moral superstructure with which to evaluate the narratives themselves is not an easy thing to do. Yet these difficulties are no excuse for failing to take up the challenge—not only ought the problem be faced squarely, it should be welcomed. Facing it puts moral evaluations out in the open, where they belong, to be debated and altered as our understandings of the ethical universe changes. The result will be frank discussions of issues of essential importance to our polity, and justice can only improve as a consequence.