ARTICLES

EXPRESSIVE THEORIES OF LAW: A SKEPTICAL OVERVIEW

MATTHEW D. ADLER[†]

INTRODUCTION	1364
I. WHAT IS AN EXPRESSIVE THEORY OF LAW?	1376
A. <i>"Expressivism," Metaethics, and Moral Philosophy</i>	1378
B. <i>Speaker's Meaning and Sentence Meaning</i>	1384
C. <i>Prescriptive Meaning and Further Meaning</i>	1396
D. <i>What Makes a Theory Genuinely Expressive?</i>	1401
E. <i>Expressive Theories of Governmental Action and</i> <i>Expressive Theories of Individual Action</i>	1408
II. EXPRESSIVE THEORIES: THE MAIN EXAMPLES	1413
A. <i>Expressive Theories of Punishment</i>	1414
B. <i>Expressive Theories of Constitutional Law</i>	1428
1. The Equal Protection Clause	1428
2. The Establishment Clause	1438
C. <i>Expressive Theories of Regulation</i>	1448
III. WHY EXPRESSIVE THEORIES ARE UNPERSUASIVE:	
A GENERAL ARGUMENT	1462
CONCLUSION	1493

[†] Assistant Professor, University of Pennsylvania Law School. B.A. Yale 1984; M. Litt. Oxford 1987; J.D. Yale 1991. Many thanks to Larry Alexander, Elizabeth Anderson, Brian Bix, Jim Chen, Claire Finkelstein, Sam Freeman, Barry Friedman, Pam Harris, Rick Hills, Heidi Hurd, Leo Katz, Seth Kreimer, Michael Moore, Mark Osiel, Stephen Perry, Rick Pildes, Eric Posner, Kenneth Simons, Steven Smith, and Alan Strudler for their helpful comments on the Article or advice about issues I confronted in writing it. All errors are my own.

INTRODUCTION

Can law be “symbolic,” “expressive,” or “meaningful”? Can it “send a message”? And, if so, should law be *evaluated* in terms of what it “symbolizes,” “expresses,” or “means”? Does the “meaning” of law bear upon its moral rightness, goodness, or legitimacy?

An “expressive” theory of law gives affirmative answers to these questions. Such a theory claims that the action of a legal official or official body can indeed be meaningful, and that the meaning thus attached to an official action is relevant to, if not determinative of, the moral status of that action. Yet this formulation is only preliminary; it needs to be further developed. For what is the meaning, here, of “meaningful”? Is law “meaningful” in the way that language is meaningful, or in some other way, such as the way that dark clouds “mean” rain?¹ And how robust a connection between the meaning of an official action (whatever precisely that entails) and the moral rightness, goodness, or legitimacy of that action must a moral theory posit, for that theory to count as “expressive”? For example, a preference-utilitarian will happily concede that if some legal actor’s saying *X* will lead to the highest degree of preference satisfaction overall, then that legal actor is under a moral obligation to say *X*. This hardly implies that the preference-utilitarian holds an “expressive” theory of law; what government says has only a contingent, and normally only a causal, connection with the satisfaction of preferences.²

In this Article, I try to clarify these conceptual issues. What, precisely, is an expressive theory of law? I then turn to the substantive question: Are expressive theories (thus clarified) persuasive? The answer, I will argue, is that they are not.

What motivates this Article is the recent popularity of expressive theories (or at least theories described as “expressive” by their proponents) within legal scholarship. I have in mind, especially, the work of Professors Richard Pildes, Dan Kahan, and Cass Sunstein. Professor Pildes has proposed an expressive theory of voting rights law in a frequently cited 1993 article about the Supreme Court’s decision in *Shaw*

¹ See PAUL GRICE, *Meaning*, in STUDIES IN THE WAY OF WORDS 213, 213-15 (1989) (distinguishing between linguistic and nonlinguistic meaning); *infra* text accompanying notes 70-72 (discussing this distinction).

² See *infra* Part I.D (distinguishing between moral impact and foundational moral relevance of speech, with specific reference to preference-utilitarianism).

*v. Reno*³ and in several follow-up articles.⁴ Pildes offers the following defense of *Shaw*, which invalidated a voting district that was drawn to increase black voting strength and that had a geographically "bizarre"⁵ shape:

Government cannot redistrict in a way that conveys the social impression that race consciousness has overridden all other, traditionally relevant redistricting values. In the Court's view, certain districts whose appearance is exceptionally "bizarre" and "irregular" suggest that impression. Plaintiffs need not establish that they suffer material harm, in the sense of vote dilution, from such a district. *Shaw* is fundamentally concerned with expressive harms: the social messages government conveys when race concerns appear to submerge all other legitimate redistricting values.⁶

Nor is Pildes merely an election law specialist. His views about voting rights are merely one component of a larger expressive theory, one that covers a wide range of constitutional provisions. These broader views are presented in a recent article, in which Pildes asserts the following:

The expressive dimension of governmental action plays a central, but underappreciated role in constitutional law. This is not a technical point with only obscure significance but a central aspect of ongoing constitutional practice with pervasive implications for the way both defenders and critics understand constitutionalism. . . .

....

[T]he [nonexpressive] view of rights as trumps fails to recognize the role of expressive harms in constitutional law. On the atomistic conception of rights, only injuries to individual interests in autonomy, dignity, self-expression, freedom of conscience, and the like, can generate constitutionally cognizable harms. But American constitutional law provides a more expansive conception of harm because it is more attuned than conventional rights theory appreciates to the social meanings of state action. Expressive harms, no less than material harms to these kind of individual interests, ground constitutional doctrine in many areas.⁷

³ 509 U.S. 630 (1993).

⁴ See Richard H. Pildes & Richard G. Niemi, *Expressive Harms, "Bizarre Districts," and Voting Rights: Evaluating Election-District Appearances After Shaw v. Reno*, 92 MICH. L. REV. 483 (1993) [hereinafter Pildes & Niemi, *Expressive Harms*]. The follow-up articles are: Richard H. Pildes, *Principled Limitations on Racial and Partisan Redistricting*, 106 YALE L.J. 2505 (1997) [hereinafter Pildes, *Principled Limitations*]; and Richard H. Pildes, *Why Rights Are Not Trumps: Social Meanings, Expressive Harms, and Constitutionalism*, 27 J. LEGAL STUD. 725, 744-47, 755-60 (1998) [hereinafter Pildes, *Why Rights Are Not Trumps*].

⁵ *Shaw*, 509 U.S. at 644.

⁶ Pildes & Niemi, *Expressive Harms*, *supra* note 4, at 526-27.

⁷ Pildes, *Why Rights Are Not Trumps*, *supra* note 4, at 760, 762; see also Richard H.

Pildes has advanced equally general claims about the expressive cast of governmental regulation—claims that will be described and criticized at some length below.⁸

Professor Kahan has not yet offered an expressive theory of law comparable in breadth to Pildes'. What Kahan *has* done, with much vigor and success, is to develop and defend an expressivist approach to a specific legal institution of great practical importance and scholarly interest—the institution of criminal punishment. In *What Do Alternative Sanctions Mean?*⁹—an article that, like Pildes' work on expressivism, has had a very high profile within the legal academy and has gained a fair measure of extra-academic attention as well¹⁰—Kahan argues that standard proposals to replace imprisonment with more cheaply administered sanctions, such as fines or community service, are politically unpalatable and normatively unattractive because they ignore the expressive dimension of punishment. Imprisonment *symbolizes* moral condemnation,¹¹ while the message communicated by these alternative sanctions is far less clearly condemnatory;¹² further, a condemnatory message is, Kahan claims, an appropriate response to criminal wrongdoing on both deterrent and retributive theories of punishment.¹³ Thus, where imprisonment is too expensive, the state should replace it with a cheaper but expressively apt sanction, specifically the sanction of "shaming":

Pildes, *Avoiding Balancing: The Role of Exclusionary Reasons in Constitutional Law*, 45 HASTINGS L.J. 711 (1994) (defending the claim, central to *Why Rights Are Not Trumps*, that constitutional law is structured by excluding certain reasons or justifications for state action).

⁸ See Richard H. Pildes & Elizabeth S. Anderson, *Slingshot Arrows at Democracy: Social Choice Theory, Value Pluralism, and Democratic Politics*, 90 COLUM. L. REV. 2121 (1990) [hereinafter Pildes & Anderson, *Slingshot Arrows*]; Richard H. Pildes & Cass R. Sunstein, *Reinventing the Regulatory State*, 62 U. CHI. L. REV. 1 (1995) [hereinafter Pildes & Sunstein, *Reinventing the Regulatory State*]; *infra* Part II.C (summarizing and criticizing the expressive theory of regulation presented in these articles).

⁹ Dan M. Kahan, *What Do Alternative Sanctions Mean?*, 63 U. CHI. L. REV. 591 (1996).

¹⁰ My Westlaw check (on April 25, 2000, in the journals and law review ("JLR") database) turned up 81 citations to Kahan's article, *What Do Alternative Sanctions Mean?*, *supra* note 9, 140 citations to Pildes' co-authored article with Niemi, *Expressive Harms*, *supra* note 4, and 145 citations to Pildes' co-authored article with Anderson, *Slingshot Arrows*, *supra* note 8. Kahan's article has, in particular, ignited a scholarly debate about the legitimacy of "shaming" penalties, *see, e.g.*, James Q. Whitman, *What Is Wrong with Inflicting Shame Sanctions?*, 107 YALE L.J. 1055 (1998), and Kahan's views on this score have been reported in the television, radio, and print media, *see* Stephen P. Garvey, *Can Shaming Punishments Educate?*, 65 U. CHI. L. REV. 733, 744 n.52 (1998).

¹¹ See Kahan, *supra* note 9, at 597-601.

¹² See *id.* at 605-30.

¹³ See *id.* at 601-05.

Punishment is *not* just a way to make offenders suffer; it is a special social convention that signifies moral condemnation. Not all modes of imposing suffering express condemnation or express it in the same way. The message of condemnation is very clear when society deprives an offender of his liberty. But when it merely fines him for the same act [or imposes a penalty of community service], the message is likely to be different: you may do what you have done, but you must pay for the privilege. . . . This mismatch between the suffering that a sanction imposes and the meaning that it has for society is what makes alternative sanctions politically unacceptable.

....

[By contrast, s]haming penalties unambiguously express condemnation and are a feasible alternative to imprisonment for many offenses.¹⁴

Kahan has pursued this defense of "shaming," and the underlying idea that the state's response to criminal wrongdoing should communicate a linguistic message (i.e., the message of condemnation), in a number of subsequent articles.¹⁵

Professor Sunstein's endorsement of expressivism is more tentative than that of Pildes or Kahan. Sunstein's original interest in this subject stems, it seems, from his work on "incommensurability."¹⁶ To say that government's choices are incommensurable means, roughly, that their comparative worth cannot be measured on a single scale—in particular, on a utilitarian scale or a cost-benefit scale.¹⁷ For exam-

¹⁴ *Id.* at 593-94. See generally *id.* at 630-52 (defending "shaming").

¹⁵ See Dan M. Kahan, *The Anatomy of Disgust in Criminal Law*, 96 MICH. L. REV. 1621, 1639-43 (1998) (reviewing WILLIAM IAN MILLER, *THE ANATOMY OF DISGUST* (1997)); Dan M. Kahan, *Punishment Incommensurability*, 1 BUFF. CRIM. L. REV. 691, 704-06 (1998); Dan M. Kahan, *Social Influence, Social Meaning, and Deterrence*, 83 VA. L. REV. 349, 382-85 (1997) [hereinafter Kahan, *Social Influence, Social Meaning, and Deterrence*]; Dan M. Kahan, *Social Meaning and the Economic Analysis of Crime*, 27 J. LEGAL STUD. 609, 615-17 (1998); cf. Dan M. Kahan, *The Secret Ambition of Deterrence*, 113 HARV. L. REV. 413, 419-35 (1999) (analyzing deterrence and expressivism as two contrasting idioms of legal discourse).

¹⁶ See Cass R. Sunstein, *Incommensurability and Valuation in Law*, 92 MICH. L. REV. 779, 820-24 (1994) [hereinafter Sunstein, *Incommensurability and Valuation*] (discussing the expressive function of law within the context of a work on incommensurability); see also Cass R. Sunstein, *Incommensurability and Kinds of Valuation: Some Applications in Law* [hereinafter Sunstein, *Incommensurability and Kinds of Valuation*], in INCOMMENSURABILITY, INCOMPARABILITY, AND PRACTICAL REASON 234, 244-45 (Ruth Chang ed., 1997) [hereinafter INCOMMENSURABILITY] (same).

¹⁷ See generally Matthew Adler, *Law and Incommensurability: Introduction*, 146 U. PA. L. REV. 1169 (1998) [hereinafter Adler, *Introduction*] (elaborating different senses of "incommensurability"); Symposium, *Law and Incommensurability*, 146 U. PA. L. REV. 1169 (1998); Matthew Adler, *Incommensurability and Cost-Benefit Analysis*, 146 U. PA. L. REV. 1371, 1383 n.47 (1998) [hereinafter Adler, *Cost-Benefit Analysis*] (citing philo-

ple: How can we compare, in dollar terms, the choice between preserving a pristine wilderness area and opening it to commercial development? How can we compare, in terms of "utils," the choice between a free market in babies and a prohibition on surrogacy and baby-selling? One reason why governmental choices of this kind might be incommensurable on a dollar or utilitarian scale—if indeed they are—is that such scales ignore the expressive considerations bearing upon these choices. This is what Sunstein proposed:

A society might identify the kind of valuation to which it is committed and insist on that kind, even if the consequences of the insistence are obscure or unknown. A society might, for example, insist on an antidiscrimination law for expressive reasons even if it does not know whether the law actually helps members of minority groups. A society might protect endangered species partly because it believes that the protection makes best sense of its self-understanding, by expressing an appropriate valuation of what it means for one species to eliminate another. A society might endorse or reject capital punishment because it wants to express a certain understanding of the appropriate course of action after one person has taken the life of another.¹⁸

Sunstein advanced a similar claim in a joint article with Professor Pildes, defending an expressive theory of regulation.¹⁹ More recently, in a 1996 piece entitled *On the Expressive Function of Law*, Sunstein backed away from the view that it is intrinsically important for government to express certain meanings, quite apart from the consequences of such expression:

[S]ome people appear to think that consequences are barely relevant, and that it is intrinsically problematic to "say," through law, that environmental amenities are ordinary goods with appropriate prices. Is this a good objection to emissions trading programs if (as we might suppose) such programs can save billions of dollars in return for the same degree of environmental protection? I do not believe [so].²⁰

On the other hand, Sunstein has continued to argue for the proposition that law's meaning can have significant, *causal* consequences, par-

sophical literature on incommensurability).

¹⁸ Sunstein, *Incommensurability and Valuation*, *supra* note 16, at 823.

¹⁹ See Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 64-72.

²⁰ Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021, 2046 (1996) [hereinafter Sunstein, *On the Expressive Function of Law*]. There is some skepticism about the intrinsic importance of expression in Sunstein's earlier article on incommensurability, see Sunstein, *Incommensurability and Valuation*, *supra* note 16, at 824, although not, as far as I can tell, in the co-authored article with Professor Pildes, see *infra* text accompanying note 294 (quoting Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 66).

ticularly in shaping social norms, and that legal officials must therefore attend to the statements communicated by the actions they perform.²¹

The work of Professors Pildes, Kahan, and Sunstein, just described, has given renewed salience and currency to expressive theories of law.²² But it bears emphasis that their scholarship is simply the most recent contribution to a much older and larger body of scholarly writing about the symbolic cast of legal decisions. For example, students of the criminal law have long debated the expressive dimension of punishment. The famous legal philosopher Joel Feinberg, in a 1965 article entitled *The Expressive Function of Punishment*,²³ rejected the then standard definition of punishment as "the infliction of hard treatment by an authority on a person for his prior failing in some respect,"²⁴ and asserted by contrast that punishment was essentially expressive—that it necessarily had a "symbolic significance largely missing from other kinds of penalties."²⁵ Specifically, Feinberg claimed: "[P]unishment is a conventional device for the expression of attitudes of resentment and indignation, and of judgments of disapproval and reprobation, on the part either of the punishing authority himself or of those 'in whose name' the punishment is inflicted."²⁶ Feinberg's article touched off a still-flourishing debate within criminal law scholarship, prompting rebuttals by (among others) C.L. Ten, Michael Moore, and Michael Davis,²⁷ and defenses by (among others) Robert

²¹ See Sunstein, *On the Expressive Function of Law*, *supra* note 20, at 2029-44; Cass R. Sunstein, *Social Norms and Social Roles*, 96 COLUM. L. REV. 903, 964-65 (1996) [hereinafter Sunstein, *Social Norms and Social Roles*].

²² In a broader sense, the same is true of Larry Lessig's article, *The Regulation of Social Meaning*, 62 U. CHI. L. REV. 943 (1995). However, because Lessig is focused not on the meaning of governmental action, but on the effect of governmental action (meaningful or not) on the "social meaning" of individual behavior, I do not count him as an expressivist. See *infra* Part I.E (distinguishing between expressivism about individuals and expressivism about governments); *infra* notes 371-74 and accompanying text (further discussing Lessig's work).

²³ Joel Feinberg, *The Expressive Function of Punishment*, 49 MONIST 397 (1965), reprinted in JOEL FEINBERG, *DOING AND DESERVING* 95 (1970).

²⁴ *Id.* at 95 (citing earlier work by A.G.N. Flew, S.I. Benn, and H.L.A. Hart).

²⁵ *Id.* at 98.

²⁶ *Id.*

²⁷ See C.L. TEN, CRIME, GUILT, AND PUNISHMENT 41-46 (1987); Michael Davis, *Punishment as Language: Misleading Analogy for Desert Theorists*, 10 LAW & PHIL. 311 (1991); Michael S. Moore, *The Moral Worth of Retribution*, in RESPONSIBILITY, CHARACTER, AND THE EMOTIONS 179, 181 (Ferdinand Schoeman ed., 1987); see also H.L.A. HART, LAW, LIBERTY AND MORALITY 60-69 (1963) (criticizing James Fitzjames Stephen's expressive theory of punishment).

Nozick, Jean Hampton, Igor Primoratz, Anthony Duff,²⁸ and, now, Professor Kahan. Kahan quite happily admits that the conceptualization of punishment as entailing the communication of a condemnatory message is not a new one.²⁹

An equally well-established line of expressivist scholarship is that concerning the Equal Protection Clause, particularly the constitutional status of race and racial discrimination. This scholarship was animated by the expressive account of school segregation offered by the Supreme Court itself in *Brown v. Board of Education*.³⁰ The Court relied quite centrally on the claim that the institution of segregation communicated a harmful message: "[T]he policy of separating the races is usually interpreted as denoting the inferiority of the negro group [and this] sense of inferiority affects the motivation of a child to learn."³¹ The idea endorsed by the Court here—that segregation stigmatized black children, that it sent an (incorrect!) message about their appropriate status—was then developed by constitutional scholars, most prominently by Paul Brest in his 1976 article about discrimination:

[One] rationale for the antidiscrimination principle [embodied by the Equal Protection Clause] is the prevention of harms which may result from race-dependent decisions. Often, the most obvious harm is the denial of the opportunity to secure a desired benefit—a job, a night's lodging at a motel, a vote. But this does not completely describe the consequences of race-dependent decision making. Decisions based on assumptions of intrinsic worth and selective indifference inflict psychological injury by stigmatizing their victims as inferior. Moreover, because acts of discrimination tend to occur in pervasive patterns, their victims suffer especially frustrating, cumulative and debilitating injuries.³²

²⁸ See R.A. DUFF, TRIALS AND PUNISHMENTS 233-66 (1986); ROBERT NOZICK, PHILOSOPHICAL EXPLANATIONS 363-97 (1981); Jean Hampton, *Correcting Harms Versus Righting Wrongs: The Goal of Retribution*, 39 UCLA L. REV. 1659 (1992); Igor Primoratz, *Punishment as Language*, 64 PHIL. 187 (1989). Other entries in the debate about expressive theories of punishment are cited by Kahan, *supra* note 9, at 594-96 nn.7-19 and accompanying text.

²⁹ See Kahan, *supra* note 9, at 594-97.

³⁰ 347 U.S. 483 (1954).

³¹ *Id.* at 494 (internal quotation omitted).

³² Paul Brest, *The Supreme Court, 1975 Term—Foreword: In Defense of the Antidiscrimination Principle*, 90 HARV. L. REV. 1, 8 (1976). For an earlier, expressivist account of the racial component of the Equal Protection Clause, see Charles L. Black, Jr., *The Lawfulness of the Segregation Decisions*, 69 YALE L.J. 421 (1960). Other prominent scholarly accounts of this kind include KENNETH L. KARST, *BELONGING TO AMERICA: EQUAL CITIZENSHIP AND THE CONSTITUTION* 15-27 (1989) [hereinafter KARST, *BELONGING TO AMERICA*]; Kenneth L. Karst, *The Supreme Court, 1976 Term—Foreword: Equal Citizenship*

This expressive rationale for the antidiscrimination principle has been less controversial than the expressive account of punishment offered by theorists such as Feinberg and Kahan. There appears to be widespread agreement among constitutional scholars that race discrimination is both meaningful and wrongful in virtue of what it means—in short, that an expressive theory is at least one component of a complete theory of the Equal Protection Clause.³³ Andrew Koppelman recently produced a book-length treatment refining this view and extending it to discrimination based on gender and sexual orientation, and to private as well as governmental actions.³⁴

Thus far, I have described the considerable and resurgent role of expressive theories within legal scholarship. But it bears mention that expressivism has also figured significantly within legal doctrines, particularly constitutional doctrines. One important example is equal protection. The expressivism of *Brown v. Board of Education* had its antecedents in the nineteenth-century case of *Strauder v. West Virginia*³⁵—which characterized a state law excluding blacks from juries in expressive terms, as constituting “an assertion of [black] inferiority”³⁶—and has been carried forward to modern equal protection case law. For example, the current Court has sought to justify the “suspect” status of all racially discriminatory statutes and programs (including affirmative action programs) by relying upon a stigma theory of discrimination. In *City of Richmond v. J.A. Croson Co.*, for example, the Court stated that: “[c]lassifications based upon race carry a danger of stigmatic harm. Unless they are strictly reserved for remedial settings, they may in fact promote notions of racial inferiority and lead to a politics of

Under the Fourteenth Amendment, 91 HARV. L. REV. 1 (1977) [hereinafter Karst, *Foreword*]; Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317 (1987); and Cass R. Sunstein, *The Anticaste Principle*, 92 MICH. L. REV. 2410 (1994). See generally ANDREW KOPPELMAN, *ANTIDISCRIMINATION LAW AND SOCIAL EQUALITY* 57-76 (1996) (citing and summarizing “stigma” theorists).

³³ See David A. Strauss, *Discriminatory Intent and the Taming of Brown*, 56 U. CHI. L. REV. 935, 941-43 nn.15, 18 & 24 (1989) (citing scholars who support the stigma account of the Equal Protection Clause, or who hold related views to the effect that the Clause prohibits “subordination” or “second-class citizenship”).

³⁴ See KOPPELMAN, *supra* note 32. Others who have extended the theory beyond the racial component of the Equal Protection Clause include KARST, *BELONGING TO AMERICA*, *supra* note 32; and Ruth Colker, *Anti-Subordination Above All: Sex, Race, and Equal Protection*, 61 N.Y.U. L. REV. 1003 (1986). It is more accurate to speak of the stigma theory of the “racial component” of the Clause, rather than the stigma theory of “race discrimination,” since some stigma theorists think that nondiscriminatory laws can stigmatize blacks and thereby be unconstitutional. See *infra* note 226.

³⁵ 100 U.S. 303 (1879).

³⁶ *Id.* at 308; see also *Plessy v. Ferguson*, 163 U.S. 537, 562 (1896) (Harlan, J., dissenting) (depicting segregation as placing a “badge of servitude” upon blacks).

racial hostility."³⁷

Another area of legal doctrine in which expressivism has been particularly influential is the Establishment Clause. Justice O'Connor, in her concurring opinion in *Lynch v. Donnelly*,³⁸ argued that the Court should employ an "endorsement" test in adjudicating Establishment Clause cases. The central question in such cases, she contended, ought to be whether the challenged practice constituted an "endorsement" of religion—whether it sent a message that a particular religion, or religion in general, was officially approved:

The Establishment Clause prohibits government from making adherence to a religion relevant in any way to a person's standing in the political community. Government can run afoul of that prohibition in two principal ways. One is excessive entanglement with religious institutions. . . . The second and more direct infringement is government endorsement or disapproval of religion. Endorsement sends a message to nonadherents that they are outsiders, not full members of the political community, and an accompanying message to adherents that they are insiders, favored members of the political community. Disapproval sends the opposite message.³⁹

Since *Lynch*, the Court has repeatedly incorporated an endorsement analysis into its Establishment Clause decisions;⁴⁰ and all nine sitting justices are now on record as favoring some such analysis.⁴¹

Finally, it bears noting that there is an apparent overlap between claims about the expressive nature of law, and claims about the role of law in shaping norms. Professor Sunstein now claims that the main expressive role of law, insofar as it is *properly* expressive, is in shaping

³⁷ *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 493 (1989); see also *Adarand Constructors, Inc. v. Peña*, 515 U.S. 200, 229 (1995) (quoting statement in *Croson* describing racial classifications as stigmatic). Cases within equal protection jurisprudence outside of the affirmative action context in which the concept of stigma has been invoked by the modern Court include: *Freeman v. Pitts*, 503 U.S. 467, 485 (1992); *Crawford v. Board of Education*, 458 U.S. 527, 544 (1982); and *Swann v. Charlotte-Mecklenburg Board of Education*, 402 U.S. 1, 26 (1971). See generally Note, *Expressive Harms and Standing*, 112 HARV. L. REV. 1313 (1999) (discussing the extent to which the Supreme Court recognizes "expressive harm" as a basis for standing, in various Equal Protection Clause and Establishment Clause contexts).

³⁸ 465 U.S. 668 (1984).

³⁹ *Id.* at 687-88 (O'Connor, J., concurring).

⁴⁰ See *infra* notes 246-56 and accompanying text (citing and discussing case law).

⁴¹ See *Capitol Square Review & Advisory Bd. v. Pinette*, 515 U.S. 753, 763-69 (1995) (opinion of Scalia, J., joined by the Chief Justice and Justices Kennedy and Thomas); *id.* at 772-83 (O'Connor, J., concurring in part and concurring in the judgment, joined by Justices Souter and Breyer); *id.* at 797-815 (Stevens, J., dissenting); *id.* at 817-18 (Ginsburg, J., dissenting).

social norms: "[M]ost straightforwardly, the law's 'statement' about, for example, the impropriety of monetary exchanges may be designed to affect social norms and in that way, ultimately to affect both judgments and behavior."⁴² Similarly, Robert Cooter, a leading norms scholar, has used the term "expression" to describe the influence of law on norms:

When many people in a community internalize an obligation, it becomes a social norm. People who internalize obligations express their commitment in various ways. Economic analysis of law, which has recently turned to the study of social norms, has said little about internalization and expression. [I attempt] to build the foundations for an economic theory of expressive law. According to the expressive theory of law, the expression of social values [i.e., the creation of norms] is an important function of the courts or, possibly, the most important function of the courts.⁴³

I mention this apparent link between expressivism and norms because of the large amount of recent legal scholarship focusing specifically on norms.⁴⁴ Is it the case, as Cooter suggests, that the norm-shaping role of law is identical to its expressive role? Are the two different, but connected? Or, notwithstanding a 1998 symposium on norms and expressivism entitled *Social Norms, Social Meaning, and the Economic Analysis of Law*,⁴⁵ is there really no important connection at all?

In sum, we have a significant recent body of scholarly work, some of it quite general, defending expressive theories of law (the work of Pildes, Kahan, and Sunstein); an even larger body of recent scholar-

⁴² Sunstein, *On the Expressive Function of Law*, *supra* note 20, at 2025.

⁴³ Robert Cooter, *Expressive Law and Economics*, 27 J. LEGAL STUD. 585, 585-86 (1998).

⁴⁴ The legal scholarship includes ROBERT C. ELLICKSON, *ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES* (1991); the various articles in Symposium, *Law, Economics, & Norms*, 144 U. PA. L. REV. 1697 (1996); the various articles in Symposium, *Social Norms, Social Meaning, and the Economic Analysis of Law*, 27 J. LEGAL STUD. 537 (1998); Lisa Bernstein, *Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry*, 21 J. LEGAL STUD. 115 (1992); Robert Cooter, *Normative Failure Theory of Law*, 82 CORNELL L. REV. 947 (1997); Steven Hetcher, *Creating Safe Social Norms in a Dangerous World*, 73 S. CAL. L. REV. 1 (1999); William K. Jones, *A Theory of Social Norms*, 1994 U. ILL. L. REV. 545; Jody S. Kraus, *Legal Design and the Evolution of Commercial Norms*, 26 J. LEGAL STUD. 377 (1997); Lessig, *supra* note 22; Richard H. McAdams, *Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination*, 108 HARV. L. REV. 1003 (1995) [hereinafter McAdams, *Cooperation and Conflict*]; Richard H. McAdams, *The Origin, Development, and Regulation of Norms*, 96 MICH. L. REV. 338 (1997) [hereinafter McAdams, *Origin, Development*]; Sunstein, *Social Norms and Social Roles*, *supra* note 21. Further entries in the legal literature are cited by Hetcher, *supra*, at 3 nn.3-4.

⁴⁵ See *Social Norms, Social Meaning, and the Economic Analysis of Law*, *supra* note 44.

ship, the “norms” scholarship, which apparently overlaps with this work on expressivism; a longstanding debate in certain specific areas of legal scholarship, most obviously criminal law and constitutional law, about the fact and import of law’s meaning; and a significant role for expressivism in certain segments of judicial doctrine, particularly Equal Protection Clause and Establishment Clause doctrine. All of this makes the time ripe, if not overdue, for a general analysis and assessment of expressive theories of law—the analysis and assessment that this Article seeks to provide.

Part I of the Article does the initial analytic work. There, I try to clarify the concept of an expressive theory. Strictly speaking, an expressive theory claims that law is meaningful in the way that *language* is meaningful. But how is language meaningful? What is the meaning of “linguistic meaning”? These questions themselves have been the subject of long-running debate within the philosophy of language, going back to Frege, Russell, and Wittgenstein. It is a standard philosophical position that *speakers have meaning*; it is also a standard position that *sentences have meaning*.⁴⁶ The speaker’s meaning of a statement is (roughly) what the speaker intended to communicate by her utterance of it; its sentence meaning is (roughly) what the statement is conventionally uttered to communicate. I argue in Part I that expressive theories of law should be understood as sentence-meaning theories, not speaker’s-meaning theories. Legal decisions typically lack speaker’s meanings. Notwithstanding the common scholarly habit, both within and outside the literature on expressivism, of referring to the “purposes” or “intentions” or “motivations” of legal bodies,⁴⁷ there typically is no such thing (at least for multimember bodies such as courts or legislatures). On the other hand, legal decisions can and do possess sentence meanings—at a minimum, and noncontroversially, the meaning that a particular action is prescribed, proscribed, or permitted.

What, then, does an expressive (sentence-meaning) theory of law consist in? A genuine expressive theory of law must involve *more* than the claim that the sentence meaning of law has a moral impact. On virtually any moral theory, language can make a moral difference; on

⁴⁶ See *infra* text accompanying notes 79-103 (defining linguistic meaning and distinguishing between speaker’s meaning and sentence meaning).

⁴⁷ See Matthew D. Adler, *Rights Against Rules: The Moral Structure of American Constitutional Law*, 97 MICH. L. REV. 1, 82 n.278 (1998) (giving references to official “purposes,” etc., within constitutional scholarship); *infra* note 56 (giving references to official “purposes,” etc., within expressivism literature).

virtually any theory, it can matter (morally speaking) whether a governmental actor performs an action that possesses a particular sentence meaning, as opposed to an action that possesses a different meaning or is meaningless. The expressivist goes beyond attributing moral impact to linguistic meaning. Rather, and more robustly, she claims that linguistic meaning has *foundational moral relevance*—that the linguistic meaning of an action figures in the best or most perspicuous description of that action. In short, the genuine expressivist claims that linguistic meaning is an irreducible moral factor. This conception of expressivism is fleshed out and defended in Part I.

Parts II and III of the Article provide a critical assessment of expressive theories, as defined in Part I. Part II focuses on the best-developed theories within legal scholarship and judicial case law: (1) the expressive theories of punishment developed by Feinberg, Kahan, et al.; (2) expressive theories of constitutional law, specifically the two kinds that have hitherto been most influential, namely the “stigma” theory of the Equal Protection Clause and the “endorsement” theory of the Establishment Clause; and (3) the expressive theory of regulation defended by Pildes and, at one time, by Sunstein, as well as by Professor Elizabeth Anderson.⁴⁸ It is obviously beyond the scope of this Article to provide a definitive verdict on these theories, but I furnish reason to be skeptical about each and every one. Each theory involves some variant of the claim that the linguistic meaning of governmental action possesses foundational moral relevance; but in no case does the claim turn out to be persuasive or even particularly plausible.

Part III generalizes the criticism offered in Part II. The failure of the particular expressive theories hitherto developed by legal scholars—the theories of punishment, constitutional law, and regulation—does not entail that no such theory is true. We need a more general argument against expressivism, and I provide such an argument in Part III. Morality is surely plural and complex—it incorporates a wide variety of moral factors, such as overall well-being, equality, status and self-respect, deontological constraints, the factors of desert and re-

⁴⁸ This discussion is meant to be exemplary, not exhaustive. An exhaustive discussion would, for example, cover Professor Pildes’ expressive theory of voting rights, *see* sources cited *supra* note 4; his general theory of constitutional law, *see id.*; expressivism about family law, *see, e.g.*, Katharine T. Bartlett, *Re-Expressing Parenthood*, 98 YALE L.J. 293 (1988); and expressivism about tort law, *see, e.g.*, Alan Strudler, *Mass Torts and Moral Principles*, 11 LAW & PHIL. 297 (1992). These particular theories surely bear critical scrutiny, but I have sought to prevent an overly long article from becoming even longer and thus have not provided that scrutiny here.

sponsibility, and other factors that, in various ways, may seem to involve linguistic meaning—but in every case the purportedly expressive factor turns out to be nonexpressive. For example, deontological constraints prohibit actors from doing or intending harm, not from saying something; individual status and self-respect are only contingently linked to the content of governmental communication; and the factor of moral desert is, again, about actors receiving what they truly deserve, not about government's utterance of condemnatory or praise-conferring statements. The connection between the linguistic meaning of a legal official's action and what truly matters, morally speaking, about that action, is a purely contingent connection (normally a purely causal connection); and this, in turn, has large implications for determining what the linguistic or symbolic content of the action ought to be.

I. WHAT IS AN EXPRESSIVE THEORY OF LAW?

This Part clarifies the concept of an “expressive theory of law.” At the threshold, let me clarify that the focus of this Article is upon *moral theories governing the actions of legal officials*. By “expressive theory of law,” I mean an expressive *moral* theory of law—a theory of the moral criteria applicable to legal officials, such that the expressive content of official action is significant within that theory (in a way to be elucidated below). Nonmoral theories of law can, conceivably, have expressive variants.⁴⁹ For example, one could conceivably have a theory about the nature of law, such that legal enactments (or a subset thereof) are seen to be necessarily expressive in some way. Or, one could have an expressive theory about the appropriate actions of legal officials—an expressive, normative theory—that is not a moral theory. For example, one could have a normative theory to the effect that (a) legal officials are obligated by the Constitution not to make certain statements, and (b) this expressive obligation flows from the text of the Constitution, in virtue of the “plain meaning” of the text, or its original understanding, and not in virtue of moral criteria.

But the scholars who have actually proposed expressive theories of law have been mainly moral in their focus.⁵⁰ Explicitly or implicitly,

⁴⁹ I am indebted to Brian Bix and Claire Finkelstein for raising this important point and pressing me to make explicit that this Article is focused upon expressive moral theories of law.

⁵⁰ See *infra* Part II (discussing expressive theories of punishment, the Equal Protection Clause, the Establishment Clause, and regulation defended by legal scholars). As should become clear from my discussion of the scholars who have proposed expressive

they have been mainly concerned with morally evaluating legal or governmental action in light of its expressive content. Thus, my own effort, in this Article, will be to explicate what makes a moral theory of governmental action "expressive" and then to criticize moral theories of this kind.

In what way is an expressive (moral) theory of law different from, say, standard utilitarianism, or from a standard justice-based theory that enjoins lawmakers to equalize the distribution of resources? The proposition that the meaning of governmental decisions has moral import, within a standard utilitarian or justice-based theory, is true but quite banal. No one disputes that if government utters a sentence with one meaning rather than another, that utterance might affect overall well-being or the distribution of resources. Expressivists are surely making a more robust and interesting claim than that. But what, precisely, does the more robust and interesting claim consist in? This important, conceptual question has to date received very little attention in the legal literature on expressivism.

I proceed to address the question as follows. Section A distinguishes between expressive theories of law and a particular position in metaethics, the position known as metaethical "expressivism." Section B distinguishes between two kinds of linguistic meaning—sentence meaning and speaker's meaning—and argues that expressive theories of law are most plausibly understood as sentence-meaning theories. Section C differentiates between the various kinds of sentence meanings that might be attached to the actions of legal officials: between the prescriptive meaning that their actions uncontroversially possess and the further, nonprescriptive meaning that is typically posited by expressive theories. Section D describes the special role that the sen-

theories of the Equal Protection and Establishment Clauses, these scholars *do* (at least implicitly) take the relevant constitutional clauses to incorporate moral norms, because their arguments are straightforwardly moral—namely, that stigmatic governmental utterances degrade the social status and therewith the self-respect of those stigmatized, violating the moral norm of equality. See *infra* text accompanying notes 232-34, 281-85; see also *infra* text accompanying notes 243-45 (suggesting that legal as opposed to moral requirements can readily make reference to expression, but that this fact is not particularly interesting, since legal requirements can readily make reference to a wide range of different features of actions and persons).

There are a few exceptions to my claim that expressivists within legal scholarship have generally proposed moral theories of governmental or legal action. For example, Joel Feinberg's expressive theory of punishment is centrally a conceptual theory about the nature of punitive sanctions—punitive sanctions consist in part of a condemnatory statement, or so Feinberg claims—and not a moral theory specifying the conditions under which government should or should not condemn criminal wrongdoing. See *infra* text accompanying notes 162-65 (discussing Feinberg's theory).

tence meaning of official action plays in a genuine, expressive theory of law: in such a theory, the sentence meaning of an official's action has foundational moral relevance, and cannot be reduced to another, more basic factor. Finally, Section E distinguishes between expressive theories of law and expressive theories of individual action. There is a temptation to conflate the two, but in fact expressivism about individual action does not entail expressivism about law, nor vice versa.

A. "*Expressivism, Metaethics, and Moral Philosophy*"

"Expressivism" is the standard name within moral philosophy for a particular metaethical position. Metaethics is the branch of moral philosophy that concerns the nature of morality. It addresses questions such as these: Do moral statements, such as "Murder is wrong" or "Social and economic advantages ought to be distributed so as to maximize the position of the least-well-off" or "That action is morally permissible," describe facts, or do they do something else? Equivalently, are moral statements true or false—are they truth-evaluable—in the way that ordinary descriptive statements about the physical world are truth-evaluable? And if the answer to this last question is affirmative, are moral facts just the same as other facts, such as physical facts, or are moral facts somehow different?⁵¹

The position known as metaethical "expressivism" has the following features. First, the metaethical expressivist is a *noncognitivist*. She thinks that moral statements are not truth-evaluable; they do not describe facts; they are not properly the subject of belief.⁵² Second, the

⁵¹ For a comprehensive yet succinct overview of contemporary metaethics, see STEPHEN DARWALL ET AL., *Toward Fin de Siecle Ethics: Some Trends, in MORAL DISCOURSE AND PRACTICE: SOME PHILOSOPHICAL APPROACHES* 3 (1997). Other good overviews are DAVID O. BRINK, *MORAL REALISM AND THE FOUNDATIONS OF ETHICS* (1989); DAVID McNAUGHTON, *MORAL VISION: AN INTRODUCTION TO ETHICS* (1988); and MICHAEL SMITH, *THE MORAL PROBLEM* (1994).

⁵² See DARWALL ET AL., *supra* note 51, at 15-19 (characterizing expressivism as a particular variant of noncognitivism); BRINK, *supra* note 51, at 19 ("[The] noncognitivist . . . denies that there are moral facts or true moral propositions or, as a result, any moral knowledge.").

More precisely, I use "cognitivism" here to mean the view that (1) moral statements are truth-evaluable, and sometimes true, in virtue of being (2) strictly descriptive or assertoric statements. The point that moral statements are "sometimes true" is needed to rule out the view that moral statements are genuinely truth-evaluable, but always false—a so-called "error theory" of morality. The second qualification, about moral statements being descriptive or assertoric, is needed because of the recent proposal that statements that are not strictly descriptive or assertoric might nonetheless be truth-evaluable in a thin or "minimal" sense of truth. On this view of truth, if cognitivism is simply defined in the traditional way—as the view that moral statements are

metaethical expressivist thinks that moral statements function to *express* certain attitudes (other than beliefs) regarding the actions or events that are putatively the subject of moral description⁵³—for example, to express the speaker's emotional reaction to these actions or events, or to express a special kind of preference for or against these, or to express a desire that the listener engage in certain behavior, or perhaps something else. The metaethical expressivist cashes moral language, not as an assertion of the existence of moral facts, but as an expression of some attitude (other than a belief) that the utterance of moral language communicates. The seminal version of metaethical expressivism—and the crudest version as well—is that set forth by A.J. Ayer more than a half-century ago:

The presence of an ethical symbol in a proposition adds nothing to its factual content. Thus if I say to someone, "You acted wrongly in stealing that money," I am not stating anything more than if I had simply said, "You stole that money." In adding that this action is wrong I am not making any further statement about it. I am simply evincing my moral disapproval of it. It is as if I had said "You stole that money," in a peculiar tone of horror, or written it with the addition of some special exclamation marks.⁵⁴

Considerably more refined and plausible versions of expressivism have been presented by R.M. Hare and, most recently, by Simon Blackburn and Allan Gibbard.⁵⁵

truth-evaluable—then expressivism (the view that moral statements function to express non-belief attitudes rather than to describe) could be a kind of cognitivism. For a discussion of the connections between cognitivism, expressivism, and a minimal view of truth, see SIMON BLACKBURN, *RULING PASSIONS* 48-83 (1998); DARWALL ET AL., *supra* note 51, at 8 n.32; James Dreier, *Expressivist Embeddings and Minimalist Truth*, 83 *PHIL. STUD.* 29 (1996); and Michael Smith, *Why Expressivists About Value Should Love Minimalism About Truth*, 54 *ANALYSIS* 1 (1994).

My narrower definition of cognitivism preserves the standard view that metaethical expressivism is a kind of noncognitivism. See DARWALL ET AL., *supra* note 51, at 15-19. Further, I believe there are good arguments in favor of a (narrowly defined) cognitivist approach to morality, not just a (traditionally defined) cognitivist approach. See *infra* note 62.

⁵³ See, e.g., DARWALL ET AL., *supra* note 51, at 17 ("Most noncognitivists are *expressivists*: they explain moral language as expressing moral judgments, and explain moral judgments as something other than beliefs."); SMITH, *supra* note 51, at 16 ("Expressivists deny that moral judgments represent the world as being one way [or] another. . . . [They] rather serve to express the judger's attitudes of approval and disapproval, or perhaps their more complicated dispositions to have such attitudes." (citations omitted)).

⁵⁴ ALFRED JULIUS AYER, *LANGUAGE, TRUTH AND LOGIC* 107 (2d ed. 1948).

⁵⁵ See SIMON BLACKBURN, *How To Be an Ethical Anti-Realist*, in *ESSAYS IN QUASI-REALISM* 166 (1993); BLACKBURN, *supra* note 52, at 48-83; SIMON BLACKBURN, *SPREADING THE WORD: GROUNDINGS IN THE PHILOSOPHY OF LANGUAGE* 181-223

Why even mention all this? I mention it both because expressive theories of law might incorrectly be understood as related to metaethical expressivism and because, best construed, they are not. As for the first point: the risk of taking the two to be related stems from a number of factors, including, of course, the fact that the words for referring to these two different things, "expressive theories of law" and "metaethical expressivism," share a key term ("expressive"), but also from several other factors as well. Certain concepts that are essential to metaethical expressivism—specifically, the concept of an *attitude* that the utterance of a moral statement expresses—are also widely deployed in the literature on legal expressivism. For example, attitudes are central to Professor Pildes' version of legal expressivism. In his most recent article on expressivism in constitutional law, Pildes defines "expressive harm" as follows:

An expressive harm is one that results from the *ideas or attitudes expressed* through a governmental action rather than from the more tangible or material consequences the action brings about. . . . Public policies can violate the Constitution not only because they bring about concrete costs but because the meaning they convey expresses inappropriate respect for relevant constitutional norms.⁵⁶

Conversely, certain concepts that are central, and may be essential, to legal expressivism—specifically, the concept of a "norm"—have also played a role in the literature on metaethical expressivism. Allan Gibbard, one of the most recent and sophisticated of the metaethical expressivists, builds his entire theory around norms.⁵⁷ Finally, there is at least a faint whiff of noncognitivism in some of the literature on expressive theories of law. Professor Sunstein certainly comes close to noncognitivism when he says that "[a]ny particular characterization or

(1984) [hereinafter BLACKBURN, *SPREADING THE WORD*]; ALLAN GIBBARD, *WISE CHOICES, APT FEELINGS: A THEORY OF NORMATIVE JUDGMENT* (1990); R.M. HARE, *MORAL THINKING* (1981).

⁵⁶ Pildes, *Why Rights Are Not Trumps*, *supra* note 4, at 755 (emphasis added). For reference to attitudes within the literature on punishment-expressivism, see, for example, FEINBERG, *supra* note 23, at 98 (stating that "punishment is a conventional device for the expression of attitudes of resentment and indignation"); within the literature on stigma, see, for example, Brest, *supra* note 32, at 8 (suggesting that stigmatic decisions are "based on assumptions of intrinsic worth and selective indifference"); and within the literature on the Establishment Clause, see, for example, *Lynch v. Donnelly*, 465 U.S. 668, 690 (1984) (O'Connor, J., concurring) ("Examination of both the subjective and the objective components of the message communicated by a government action is therefore necessary to determine whether the action carries a forbidden meaning.").

⁵⁷ See GIBBARD, *supra* note 55, at 45-48 (summarizing "norm-expressivist" approach).

accounting of consequences will [not] rest . . . on some depiction of the brute facts; instead, it will be mediated by a set of (often tacit) norms determining how to describe or conceive of consequences."⁵⁸ Similarly, Sunstein and Pildes seem to espouse a kind of noncognitivism in their co-authored work when they state: "[The] expressive or symbolic dimensions of policy are central [because p]art of what policy-making does is to define, interpret, and *create* collective understandings and values. . . . Decisions today crystallize collective understandings in ways that shape the perceived meaning *and appropriate resolution* of future choices."⁵⁹

Notwithstanding these seeming intermittent denials of the existence of moral truth and moral facts by Professors Sunstein, Pildes, and other legal expressivists, and notwithstanding their apparent or real flirtation with metaethical expressivism, I suggest that the best way to achieve clarity about expressive theories of law is simply to place metaethical expressivism to one side. Expressive theories of law, of the kind described in the Introduction, are best analyzed and assessed within a *cognitivist* moral framework, which takes moral statements to be descriptive statements that assert the existence of moral facts and are genuinely truth-evaluable (and sometimes true).⁶⁰ Metaethical expressivism, which as I have said is a variant of noncognitivism, is thus mentioned here only to be ignored for the remainder of the Article.

My reasons for analyzing expressive theories of law within a cognitivist framework and specifically for assuming metaethical expressivism to be false are as follows: Noncognitivist views (including both metaethical expressivism and other noncognitivist views) are subject to a variety of objections well rehearsed in the philosophical literature. The noncognitivist must reject a basic feature of our moral discourse: that we criticize each other's moral judgments, claims, and arguments as right or wrong, true or false.⁶¹ Relatedly, she has serious difficulty

⁵⁸ Sunstein, *On the Expressive Function of Law*, *supra* note 20, at 2048.

⁵⁹ Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 70 (emphasis added).

⁶⁰ Note that the cognitivist framework for analyzing expressive theories, as I have presented it here, is a narrowly cognitivist framework. See *supra* note 52 (providing narrow definition of cognitivism).

⁶¹ See, e.g., BRINK, *supra* note 51, at 7-8 ("Although various sorts of considerations support moral realism, its intuitive appeal [and *a fortiori* that of cognitivism] derives . . . from the way it explains the point and nature of moral inquiry. . . . We think people can be morally *mistaken* and some people are morally more *perceptive* than others."). If the common-sense view of morality is taken to be the view that moral claims are truth-evaluable, then this view constitutes support only for cognitivism understood

explaining how we can properly reason from moral premises to moral conclusions using the mundane rules of logical inference that we routinely employ, since these rules apparently presuppose that the premises and conclusions are truth-evaluable.⁶² Further, the proposition that motivates noncognitivism—the proposition that moral facts, if they were to exist, could not be just the same as physical facts—is arguably correct,⁶³ but it is also surely much too hasty to conclude from this proposition that moral facts do not exist at all. A wide variety of statements are uncontroversially true or false, without those statements being just like statements of physical fact. For example, the statement “this is red” can be true or false, even though the redness of an object (unlike its physical features) has no robust causal role,⁶⁴ and even though the property of redness (unlike physical features) makes essential reference to an *observer*.⁶⁵ Moral facts, like facts about color, and unlike physical facts, might indeed lack a robust causal role, and might make essential reference to observers, but none of this entails

as truth-evaluability and not specifically for narrow cognitivism. If the common-sense view is taken to be the view that moral claims function to describe, then narrow cognitivism does gain support from common sense.

⁶² Here is a simple example: (1) “Lying is wrong”; (2) “If lying is wrong, murder is wrong”; therefore (3) “Murder is wrong.” If moral statements are descriptions, the inference from the premises of this syllogism to its conclusion is logically unimpeachable, since the premises and conclusion are (for purposes of the rules of logic) no different than ordinary descriptive statements like “Socrates is a human being,” “If Socrates is a human being, then he has 46 chromosomes,” and “Socrates has 46 chromosomes.” By contrast, despite intensive efforts on this score, see, e.g., BLACKBURN, *supra* note 52, at 68-77; BLACKBURN, SPREADING THE WORD, *supra* note 55, at 189-96; GIBBARD, *supra* note 55, at 92-102, noncognitivists have not yet satisfactorily explained how we can reason from (1) and (2) to (3). The seminal statement of this problem for expressivists and, more generally, noncognitivists—the so-called “embedding” problem—is P.T. Geach, *Assertion*, 74 PHIL. REV. 449 (1965). Modern criticisms of Blackburn’s and Gibbard’s purported solutions to the embedding problem include: Bob Hale, *Realism and Its Oppositions*, in A COMPANION TO THE PHILOSOPHY OF LANGUAGE 288-91 (Bob Hale & Crispin Wright eds., 1997); Paul Horwich, *Gibbard’s Theory of Norms*, 22 PHIL. & PUB. AFF. 67 (1993); Walter Sinnott-Armstrong, *Some Problems for Gibbard’s Norm-Expressivism*, 69 PHIL. STUD. 297 (1993); and Mark Van Roojen, *Expressivism and Irrationality*, 105 PHIL. REV. 311 (1996). I take these criticisms to be support for narrow cognitivism, in that they show Blackburn’s and Gibbard’s non-descriptivist analyses of moral language not to have solved the embedding problem.

⁶³ But see DARWALL ET AL., *supra* note 51, at 24-30 (describing cognitivists who take moral facts to be more or less similar to physical and other scientific facts).

⁶⁴ Cf. Hale, *supra* note 62, at 296 (discussing the “Wide Cosmological Role” that certain facts or properties play, and that arguably justifies realism about them).

⁶⁵ The *locus classicus* for understanding moral properties as a kind of “secondary quality,” epitomized by color properties, is John McDowell, *Values and Secondary Qualities*, in MORALITY AND OBJECTIVITY 110 (Ted Honderich ed., 1985). See generally DARWALL ET AL., *supra* note 51, at 19-24 (summarizing McDowell’s view); L.W. SUMNER, WELFARE, HAPPINESS AND ETHICS 27-34 (1996) (same).

that moral facts are illusory.

It is beyond the scope of this Article to make a case for moral cognitivism or even to describe in any detail the case that others have made and that I have just sketched out in the most cursory way. Let me simply note that cognitivism is now probably the dominant position within moral philosophy and that metaethical expressivism is probably a minority view.⁶⁶ Further, I am inclined to think that the legal scholars who have written about the expressive function of law are at bottom cognitivists, despite their intermittent, apparent skepticism about moral truth. The central purpose of these scholars is morally evaluative and critical: it is to evaluate legal decisions and institutions as morally good or bad, right or wrong, positive or negative, and to criticize the decisions and institutions that are found to be bad, wrong, or negative. If noncognitivism about morality is correct, then the morally evaluative and critical statements that are pervasive in the literature on legal expressivism and that take a descriptive, truth-evaluable form must be reformulated. For example, Pildes' claim that bizarre-shaped voting districts "expressively endors[e] an *impermissible* view of the constitutionally regulated role of race in public life"⁶⁷ must be reformulated as the claim that these districts "expressively endorse a view about the constitutionally regulated role of race in public life *that I do not approve*," or something like that. Similarly, the statement by Pildes and Sunstein that "CBA [cost-benefit analysis] . . . cannot *adequately* capture all the expressive dimensions of policy choices"⁶⁸ must be translated into a form that does not take normative language like "adequate" to be descriptive and truth-evaluable, namely: "CBA cannot capture all the expressive dimensions of policy choices *in a way that we want expressive dimensions to be captured*." It is hard to think that Pildes, Sunstein, or other expressivists would be comfortable with such noncognitivist reformulations, which dramatically weaken the critical and evaluative force of their scholarly texts.

So to repeat: this Article will analyze and assess expressive theories of law within a cognitivist moral framework. Noncognitivism, and specifically metaethical expressivism, will henceforth be assumed to be incorrect. But how is it possible to be a moral cognitivist and, at the same time, to think that law has an expressive function? How is it possible to be a moral cognitivist and, at the same time, a legal expressiv-

⁶⁶ See sources cited *supra* note 51.

⁶⁷ Pildes, *Why Rights Are Not Trumps*, *supra* note 4, at 760 (emphasis added).

⁶⁸ Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 70 (emphasis added).

ist? Very easily. Someone who holds these views in combination believes not only that (1) moral statements are true or false and describe moral facts (cognitivism), but also that (2) it is *truly* morally good or bad, right or wrong, positive or negative, for legal decisions to possess certain meanings (legal expressivism). If I say, "race-discrimination sends a message that blacks are inferior, and for government to send that message is truly morally wrong," I have committed myself both to cognitivism and to some kind of expressive theory of law. If I say "government's use of cost-benefit analysis to assess measures that save human lives communicates contempt for human life, and it is truly wrong for government to communicate that message," I have again committed myself both to cognitivism and to some kind of expressive theory of law. The cognitivist who does that is no weirder than the utilitarian cognitivist who says that "race discrimination reduces overall well-being, and for government to make a decision that reduces overall well-being is truly morally wrong." Expressive theories of law do not presuppose noncognitivism or metaethical expressivism any more than utilitarianism does.

B. *Speaker's Meaning and Sentence Meaning*

An expressive theory of law, in short, claims that the actions of legal officials are truly better or worse, right or wrong, legitimate or illegitimate, in virtue of what the actions mean. But what kind of "meaning" is posited here? I suggest that expressive theories are plausible, if at all, as *sentence-meaning* rather than *speaker's-meaning* theories, where the "sentence meaning" of an official action⁶⁹ is its meaning pursuant to the very same kind of rules and conventions that assign meaning to a well-formed English sentence.

Let us step back a second. "Meaning" can refer either to nonlinguistic meaning or to linguistic meaning.⁷⁰ H.P. Grice, in his seminal article on this topic, gave the following examples of nonlinguistic meaning: "Those spots mean (meant) measles"; "Those spots didn't mean anything to me, but to the doctor they meant measles"; "The recent budget means that we shall have a hard year."⁷¹ Roughly speaking, an action, event, or state of affairs has nonlinguistic meaning if it provides evidence of something; that thing is what is "meant" by the

⁶⁹ I will, throughout the Article, use the terms "legal decision," "action of a legal official," "governmental action," and similar terms interchangeably.

⁷⁰ See GRICE, *supra* note 1, at 213-15.

⁷¹ *Id.* at 213.

action, event, or state, where "meant" is synonymous with "evidenced," "indicated," or "signaled."⁷² As is clear from this definition, and as Grice's examples show, many entities besides linguistic utterances or sentences of a language can have nonlinguistic meaning. The actions of legal officials, in particular, can have a host of different Gricean nonlinguistic meanings. They can nonlinguistically *mean* (evidence, signal, or indicate) all the different things for which their performance provides evidence. For example, the enactment of a particular statute might nonlinguistically mean that constituents have certain preferences, that the Republicans will win the next election, that special interest groups have been relatively active or inactive in their lobbying, that the President made a strategic error, that most Americans are racist, and so on.

But this evidentiary feature of statutes and other legal decisions is not, I take it, what expressivists intend to refer to when they describe the actions of legal officials as being "meaningful." Rather, they claim (or at least many of them do) that official actions are meaningful in the way that *language is meaningful*. Punishing crime *C* is meaningful and, in particular, expresses condemnation, in the very same way that the utterance, "crime *C* is wrong" is meaningful and expresses condemnation, or so expressivists about punishment standardly claim.⁷³ Discriminating against black children is meaningful and, in particular,

⁷² See Heidi M. Hurd, *Sovereignty in Silence*, 99 YALE L.J. 945, 953 (1990) ("[Entities] possessing what H.P. Grice termed 'natural' [i.e., nonlinguistic] meaning . . . function as symptoms of conditions in the world to which they are causally related.").

⁷³ See, e.g., DUFF, *supra* note 28, at 235-36 ("Punishment . . . expresses condemnation: it denounces and formally disapproves the criminal's act; it disavows that act as one which is not to be tolerated or condoned. . . . [W]e can express a formal disapproval of someone's criminal wrong-doing simply by the words which constitute a criminal conviction: but the further punitive measures which are then imposed on a criminal can also express condemnation"); FEINBERG, *supra* note 23, at 98 ("[P]unishment is a conventional device for the expression of [certain attitudes and judgments] [and] has a *symbolic significance* largely missing from other kinds of penalties."); NOZICK, *supra* note 28, at 370 ("The complicated structure of the nine conditions for retribution [presented by Nozick], wherein something intentionally is produced in another with the intention that he realize why it was produced and that he realize he was intended to realize all this, fits the account of [linguistic] meaning offered by H.P. Grice. . . . Retributive punishment is an act of communicative behavior."); HAMPTON, *supra* note 28, at 1685-98 (defending an account of retribution that includes a linguistic component); KAHAN, *supra* note 9, at 594 (arguing that "punishment is usefully conceived of as a language"); PRIMORATZ, *supra* note 28, at 187, 196-98 (defining "expressionism" as the view that the "evil inflicted on the person punished is not an evil *simpliciter*, but rather the expression of an important social message—that punishment is a kind of language," and defending a variant of "expressionism" thus defined).

labels them as inferior, in the very same way that an official edict that reads, "blacks are hereby declared to be inferior" is meaningful and labels them as inferior,⁷⁴ or so expressivists about the Equal Protection Clause standardly claim.⁷⁵ Placing a crèche on city property is meaningful and, in particular, constitutes an endorsement of Christianity, in the very same way that a governmental declaration, "We hereby endorse Christianity" does, or so expressivists about the Establishment Clause standardly claim.⁷⁶ We shall briefly turn, in the Conclusion, to

⁷⁴ The term "labels" here is intentionally ambiguous between the view that to stigmatize is to invoke a descriptive linguistic convention, and the view that it is to invoke a declarative one. On this issue, see *infra* text accompanying note 223.

⁷⁵ See, e.g., KOPPELMAN, *supra* note 32, at 57-76 (describing stigma theorists of race discrimination, including Black, Brest, and Karst, and defining stigma as a "mark . . . that defines the bearer as deviant, flawed, or otherwise undesirable"); Black, *supra* note 32, at 425 ("[Segregation is intended] to stamp [blacks] with the mark of inferiority."); Brest, *supra* note 32, at 8-9 (endorsing the view that racial discrimination "stigmatizes" blacks, and explaining the concept of stigma by reference to a statement in *Strauder v. West Virginia* that discrimination is "practically a brand upon [blacks] . . . [and] an assertion of their inferiority" (internal quotations omitted)); Karst, *Foreword*, *supra* note 32, at 48, 49 ("The chief target of the equal citizenship principle [animating the equal protection clause] is the stigma of caste," which paradigmatically although not exclusively involves a "deliberate legislative choice to impose a badge of inferiority."); Lawrence, *supra* note 32, at 355-56 (proposing a test that would "evaluate governmental conduct to see if it conveys a symbolic message to which the culture attaches racial significance" and noting that "certain actions, words, or signs may take on meaning within a particular culture as a result of the collective use of those actions, words, or signs to represent or express shared but repressed [racist] attitudes"). The linguistic construal of stigma is not universal—some stigma theorists are less concerned with governmental actions that place a mark of inferiority on blacks than with actions that have a certain, negative cultural effect upon them, see *infra* notes 241-42 and accompanying text—but it is certainly quite standard, as the quotations here are meant to suggest. As David Strauss puts it: "[The stigma] approach [to equal protection] focuses . . . on the message that the [governmental] action conveys to others. Stigma in this sense is related to defamation." Strauss, *supra* note 33, at 942.

⁷⁶ See, e.g., *Lynch v. Donnelly*, 465 U.S. 668, 690 (1984) (O'Connor, J., concurring) ("The central issue in this case is whether [the city] has endorsed Christianity by its display of the crèche. To answer that question, we must examine both what [the city] intended to communicate . . . and what message the . . . display actually conveyed."); Kenneth L. Karst, *The First Amendment, the Politics of Religion and the Symbols of Government*, 27 HARV. C.R.-C.L. L. REV. 503, 517-18 (1992) ("Problems both definitional and institutional embarrass the effort to make the endorsement of religion into a generalized legal test governing the whole range of Establishment Clause issues. . . . [These] problems diminish, however, when the case at hand involves an official governmental display of a religious symbol. Whatever else can be said about such a display, it communicates a message and is so intended. . . . When the symbols of government are religious symbols, then, the endorsement test seems [an appropriate] doctrinal formula . . ."); Gary C. Leedes, *Rediscovering the Link Between the Establishment Clause and the Fourteenth Amendment: The Citizenship Declaration*, 26 IND. L. REV. 469, 507 (1993) ("It violates enduring American traditions of justice if state supported religion 'sends a message [of endorsement] to nonadherents' [of the endorsed religion].") (citation

theories of law that are predicated on the nonlinguistic meaning possessed by the actions of legal officials,⁷⁷ but the focus of this Article will be upon linguistic meaning.⁷⁸

This brings me to the distinction between *speaker's meaning* and *sentence meaning*. That distinction is a distinction within the category of linguistic meaning. Speaker's meaning and sentence meaning constitute two different ways in which language is meaningful; they constitute two different ways in which words are *symbols*, and not just evidence or signals.⁷⁹ Very roughly, the speaker's meaning of a linguistic utterance is what the speaker intends by that utterance. If I say, "that man is wearing a hat," knowing that you are a novice speaker of English who understands "is wearing a hat" the same way that fluent speakers of English understand "is a postman," and if I make the utterance in order to get you to believe that the man is a postman, then the speaker's meaning of my utterance is: that man is a postman. That is what I intend the utterance to mean; that is the belief that I intend to induce in you by my performance of it.⁸⁰ Similarly, if I draw

omitted)); Arnold H. Loewy, *Rethinking Government Neutrality Towards Religion Under the Establishment Clause: The Untapped Potential of Justice O'Connor's Insight*, 64 N.C. L. REV. 1049, 1051, 1069 (1986) (stating that "government cannot convey a message that anyone is inferior or superior because of his or her religion" and concluding that an endorsement test "is well suited to preventing successful government attempts, whether subtle or overt, to impose a 'badge of inferiority' on our religious minorities").

⁷⁷ See *infra* text accompanying notes 446-51. Andrew Altman's recent article, *Expressive Meaning, Race, and the Law: The Racial Gerrymandering Cases*, 5 LEGAL THEORY 75 (1999), is focused on expressivism in various nonlinguistic senses. See *id.* at 75, 77. I agree with Altman that expressivism might be thus understood, but I do not agree that extant scholarly theories described by their proponents as "expressive" are generally concerned with nonlinguistic rather than linguistic meaning. See *id.* at 75.

⁷⁸ Expressivists about regulation are also, at least arguably, concerned with linguistic meaning. See, e.g., Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 66 ("By expressive dimensions—what might be understood as cultural consequences of choice—we mean the values that a particular policy choice, in the specific context in which it is taken, will be generally understood to endorse."). Alternatively, the reference to "cultural consequences" here might be read to support the view that regulation-expressivists really are concerned with expression in a nonlinguistic sense, such as a cultural-impact sense or some other nonlinguistic sense. For more on this issue, see *infra* note 294.

⁷⁹ For general discussions of linguistic meaning and of the subcategories of speaker's meaning and sentence meaning, see BLACKBURN, *SPREADING THE WORD*, *supra* note 55, at 110-40; ROBERT M. MARTIN, *THE MEANING OF LANGUAGE* 65-95 (1987); Anita Avramides, *Intention and Convention*, in *A COMPANION TO THE PHILOSOPHY OF LANGUAGE* 60-86 (Bob Hale & Crispin Wright eds., 1997). For a discussion with specific reference to law, see Hurd, *supra* note 72.

⁸⁰ This is true on Grice's definition of speaker's meaning; it may not be true on others. See MARTIN, *supra* note 79, at 85-93 (summarizing objections to Grice's definition); *infra* note 82 (citing alternative definitions).

a picture of a man with a big shoe and show the picture to you, hoping to get you to realize that the postman is coming (we have an inside joke that postmen, because they walk so much, need big shoes), then my drawing of the picture has the speaker's meaning, "the postman is coming," and in that way is a linguistic utterance. Grice himself thought that speaker's meaning was the most basic kind of linguistic meaning, and offered the following analysis of it:

Perhaps we may sum up what is necessary for *A* to mean something by *x* [that is, to linguistically mean something by *x*] as follows. *A* must intend to induce by *x* a belief in an audience, and he must also intend his utterance to be recognized as so intended. But these intentions are not independent; the recognition is intended by *A* to play its part in inducing the belief, and if it does not do so something will have gone wrong with the fulfillment of *A*'s intentions. Moreover, *A*'s intending that the recognition should play this part implies . . . that he assumes that there is some chance that it will in fact play this part, that he does not regard it as a foregone conclusion for the belief will be induced in the audience whether or not the intention behind the utterance is recognized.⁸¹

Yet more refined and complex analyses of speaker's meaning have subsequently been offered. What they share with Grice's first and seminal analysis is some reference to the actual communicative intentions of the speaker—to what the speaker actually meant by her utterance.⁸²

⁸¹ GRICE, *supra* note 1, at 219.

⁸² See, e.g., STEPHEN B. SCHIFFER, *MEANING* (1972); JOHN R. SEARLE, *SPEECH ACTS: AN ESSAY IN THE PHILOSOPHY OF LANGUAGE* (1969); William P. Alston, *Illocutionary Acts and Linguistic Meaning*, in *FOUNDATIONS OF SPEECH ACT THEORY: PHILOSOPHICAL AND LINGUISTIC PERSPECTIVES* 29 (Savas Tsohatzidis ed., 1994) [hereinafter *FOUNDATIONS OF SPEECH ACT THEORY*]; John R. Searle, *What Is a Speech Act?*, in *THE PHILOSOPHY OF LANGUAGE* 129-30 (A.P. Martinich ed., 1985) [hereinafter *Searle, What Is a Speech Act?*]; P.F. Strawson, *Intention and Convention in Speech Acts*, 73 *PHIL. REV.* 439 (1971). See generally Avramides, *supra* note 79, at 71-78 (summarizing critiques of and refinements to Gricean definition).

Although some of these subsequent definitions of speaker's meaning tie it more closely to sentence meaning, I take it that they still leave *some* scope for an utterance to possess a speaker's meaning distinct from its sentence meaning. For example, Searle's definition emphasizes the conventional element of speech acts:

In the performance of [a speech act] the speaker intends to produce a certain effect by means of getting the hearer to recognize his intention to produce that effect, and furthermore, if he is using words literally, he intends this recognition to be achieved in virtue of the fact that the rules for using the expressions he utters associate the expressions with the production of that effect.

Searle, *What Is a Speech Act?*, *supra*, at 130. But Searle also notes that "[i]n hints, insinuations, irony, and metaphor—to mention a few examples—the speaker's utterance meaning and the sentence meaning come apart in various ways." JOHN SEARLE, *Indirect*

But whatever the importance of speaker's meaning in a general theory of language, as one kind or even the most basic kind of linguistic meaning, it will hardly work as the foundation for an expressive theory of law. The actions of legal officials typically do not possess speaker's meanings. More precisely, the actions of multimember legal bodies—the enactment of statutes by legislatures; the issuance of opinions by multimember courts; the promulgation of regulations by multimember agencies—typically do not possess speaker's meanings. Although such actions *are* partially constituted by words and language—specifically, by the official text that a majority of the multimember body has voted to approve—there is typically no actual intention behind this text that would give purchase to a speaker's-meaning theory of law.⁸³

To begin, legislatures, courts, agencies, and other legal institutions do not possess mental states, independent of the mental states of the persons that make up these institutions.⁸⁴ If every member of the legislature intends M_1 and solely M_1 by her enactment of the statute, then (at best) the legislature itself intends M_1 and solely M_1 ; it cannot intend M_2 if no member herself intends that. We might say, metaphorically, that the legislature “intends M_2 ,” but the term “intends” here will not refer to an actual mental state of anyone—of anything that has a mind—and will therefore not signify the kind of intention that lends speaker's meaning to actions. For example, we could say metaphorically that the legislature in this case “intends M_2 ”—and therefore, that the enactment of the statute “means M_2 ”—because that enactment leads the population to believe that M_2 is the case or to be-

Speech Acts, in *EXPRESSION AND MEANING* 30 (1978) [hereinafter SEARLE, *Indirect Speech Acts*]. Even if it were to turn out that the best analysis of linguistic meaning leaves no scope for an utterance to possess a speaker's meaning distinct from its sentence meaning—and I am skeptical that that will turn out to be the case—that would not undermine my thesis in this Section, which is that the kind of linguistic meaning relevant for expressive theories of law is sentence meaning. See *infra* text accompanying notes 83-103.

⁸³ See generally Hurd, *supra* note 72, at 968-76 (arguing that statutes do not possess speaker's meaning, specifically because legislatures are not “speakers”). As Hurd puts it:

The first challenge facing any attempt to describe statutes as speech acts is the well-known difficulty of conceiving of the legislature as having any intentions at all, let alone the kinds of communicative intentions which Grice and others have demonstrated are necessary. . . . [T]here can be no communication unless there is a speaker performing a speech act with an utterance; whether there is such a speaker and speech act wholly depends upon whether the requisite intentions are possessed.

Id. at 968-69.

⁸⁴ See *id.* at 969-70.

lieve that the legislature intends M_2 . But such "meaning" is no more linguistic than the case of clouds meaning rain. Genuine communication presupposes a genuine speaker or—a point we will get to in a moment—genuine rules, conventions, or practices for attaching significance to actions,⁸⁵ both of which are lacking here. If the arguments for, or implications of, a linguistic-meaning theory of law are plausibly different from the arguments for, or implications of, a mere signaling theory (or some other nonlinguistic theory)—and I think the arguments and implications *are* plausibly different—it is a significant conceptual mistake to lump together the case of the legislature "meaning M_2 " notwithstanding individual members solely intending M_1 with the case of a genuine speaker's meaning.

So if the action of a multimember legal body is to possess a genuine speaker's meaning, the collective intention that confers such meaning upon that action must be some function of the individual intentions of the members. Either the collective intention is (1) an intention that is shared by all the members, or least by those who approved the action, or it is (2) some other function of the members' individual intentions.

As for the first option: the members of a body that issues a legal text typically do not share any intentions with respect to that text, beyond the intention to write the text into law.⁸⁶ If the legislature approves a racially discriminatory law, all we can typically say is that the legislators who voted for the law shared the intention to enact it; we cannot, typically, ascribe some further intention to them, or at least

⁸⁵ See *infra* text accompanying notes 92-103.

⁸⁶ See Hurd, *supra* note 72, at 971-73 (detailing difficulties with the view that legislative intention can be understood "as a summation of the intentions shared by a majority of [the legislature's] members"). For similar skepticism about legislators sharing intentions with respect to a statute, beyond the intention to enact the statute into law, see Jeremy Waldron, *Legislators' Intentions and Unintentional Legislation*, in *LAW AND INTERPRETATION: ESSAYS IN LEGAL PHILOSOPHY* 329 (Andrei Marmor ed., 1995) [hereinafter *LAW AND INTERPRETATION*]; Kenneth A. Shepsle, *Congress Is a "They," Not an "It": Legislative Intent as Oxymoron*, 12 INT'L REV. L. & ECON. 239 (1992). But see Larry Alexander, *All or Nothing at All? The Intentions of Authorities and the Authority of Intentions*, in *LAW AND INTERPRETATION*, *supra*, at 357 (arguing that legal texts should be interpreted as their drafters intended, insofar as legal texts function to provide an authoritative determination of what ought to be done); Daniel A. Farber & Philip P. Frickey, *Legislative Intent and Public Choice*, 74 VA. L. REV. 423 (1988) (arguing that "legislative intent" is coherent and should play an important role in statutory interpretation). Skepticism about legislators sharing intentions is, in part, what motivates the "textualist" approach to statutory interpretation. This approach is defended most vigorously by Justice Scalia. See William N. Eskridge, Jr., *The New Textualism*, 37 UCLA L. REV. 621, 640-56 (1990).

not a further intention to communicate anything. (Remember how detailed a Gricean communicative intention is: "A must intend to induce by α a belief in an audience, and he must also intend his utterance to be recognized as so intended. [Moreover, this very recognition must be] intended by A to play its part in inducing the belief . . ."⁸⁷ This degree of detail seems to make it particularly unlikely that the individual members of multimember legal bodies routinely share communicative intentions.) Now, perhaps my claim here is a bit too strong: if a law has sentence meaning M , then perhaps we can say that the legislators who voted for the law shared both the intention to write that text into law and the intention to communicate M . For example, if race discrimination is a conventional device by which to communicate the inferiority of blacks, perhaps we can say that the legislators who approve a discriminatory statute intend to communicate black inferiority. This approach, however, makes our speaker's-meaning theory of law parasitic upon a sentence-meaning theory.

Option (2) is too counterintuitive to consider at length. Imagine that one legislator who votes for a statute has an illicit communicative intention, but no one else in the enacting coalition does, and that the statute would have been approved even if the one legislator had voted differently. Would it make sense, on an expressive theory of constitutional law, to strike down the *legislature's* action—the enactment of the statute—as unconstitutional? Slightly more plausible is the view that the statute is unconstitutional if the lone legislator was a but-for cause of enactment. But this view runs afoul of the objection that various sorts of illicit communicative intentions can cause the enactment of statutes without conferring linguistic meaning upon the statutes. (For example, if a large donor says something illicit to the majority whip, which in turn leads the whip to work feverishly for the statute's enactment—because the whip is scared of electoral defeat, not because he shares the donor's views—then the resulting statute is caused by the donor's statement but hardly means what the donor said.) The action of enacting a statute is a joint action by legislators; it is a necessary condition, for that action to occur, that a majority of legislators vote for the statute; and thus it would also seem to be a necessary condition for the joint action to be attributed a particular speaker's meaning (within a moral theory of legislation) that *this* was the meaning of at least a majority of the legislature.⁸⁸

⁸⁷ GRICE, *supra* note 1, at 219.

⁸⁸ Even if I am wrong about this, there is the further point that the best evidence of speaker's meaning accessible to the audience for legislation will be its sentence mean-

At this point, the speaker's-meaning theorist might return to option (1), dig in her heels, and insist that the expressive criteria of her theory do refer to the shared communicative intentions of all the members (or all the approving members) of multimember legal bodies—even though such intentions are not *typically* shared. On this view, the expressive constraint against racial stigma in the Equal Protection Clause is triggered by a regulation if and only if all the agency members, or at least all the members who voted to promulgate the regulation, actually intended to communicate the inferiority of blacks; the expressive constraint against the endorsement of religion in the Establishment Clause is triggered by a statute if and only if all the legislators, or at least all the legislators who voted to enact the statute, actually intended to endorse religion; and so on. This would mean, for example, that if the city council approves the placement of a crèche on city property, together with the issuance of a declaration that "Christianity is the city's official religion," but a majority of the individual council members do *not* (as it happens) intend to endorse Christianity, then no Establishment Clause violation has ensued. That is a possible position, but not one that, as far as I am aware, any legal expressivist has ever advocated.

Further, the position, albeit possible, is implausible. It is implausible because the best evidence of speaker's meaning accessible to whatever *audience* is plausibly stipulated by the expressive theorist (whether that audience is composed of the citizenry, or of lower-level legal officials, or both) will simply be the sentence meaning of legal actions.⁸⁹ If the council enacts the crèche ordinance, the audience will conclude that the council members intended to endorse Christianity. Conversely, if the council enacts an ordinance placing a car in the city park, and a majority of the council intended to endorse Christianity by that action, the attempted communication will fail, because the citizenry (or lower-level officials) will fail to realize what the council members intended. So why insist, despite this, that an expressive theory of law refers to speaker's meanings, not sentence meanings?

I have taken considerable space here to rebut a speaker's-meaning theory of law, given the common scholarly ascription of "attitudes,"

ing. I make this point, immediately below, with respect to variant (1) of the speaker's-meaning theory. The same point holds true of variant (2).

⁸⁹ This point, as well as the need for a general theory to encompass both multimember legal bodies and individual legal officials, explains why the best expressive theory covering the actions of such individuals is a sentence-meaning theory—notwithstanding the fact that these actions, by contrast with those of multimember bodies, *may* routinely have speaker's meanings apart from their sentence meanings.

"purposes," "motives," and "intentions" to legal bodies or officials—often encapsulated in claims that some body's or official's action expresses an impermissible attitude, purpose, motive, or intention.⁹⁰ Professor Pildes, who certainly engages in this kind of talk himself, ultimately concludes that the ascriptions are objective, not subjective—that they do not refer to the actual mental states of the members of multimember legal bodies or the actual mental states of other legal officials. "[I]t is the social meaning state action conveys, rather than subjective intent, that is central to applying the structural conception [i.e., Pildes' conception] of rights."⁹¹ For the reasons that I have just articulated at some length, Pildes is correct in ultimately rejecting a speaker's-meaning theory of law.

We are left, then, with sentence meaning. What is that? Roughly, sentence meaning is the kind of linguistic meaning attached to a well-formed sentence, independent of the communicative intentions (or other mental states) that happen to be held by those who utter the sentence. "Grass is green" is sentence-meaningful even if the utterer intends to communicate something different, or possesses no communicative intention at all, or indeed possesses no intention at all (say, if he is an automaton who mechanically produces the string of symbols constituting this sentence).⁹² Sentence meaning is therefore a much more attractive basis than speaker's meaning for an expressive theory of law. Legal decisions surely can and do possess sentence

⁹⁰ See *supra* note 56 (citing such ascriptions).

⁹¹ Pildes, *Why Rights Are Not Trumps*, *supra* note 4, at 752-53.

⁹² See Hurd, *supra* note 72, at 966 ("[S]tatutes [insofar as they possess sentence meaning] might be like the often-hypothesized novel typed by random chance by the thirteen-thousandth monkey chained to a typewriter: meaningful and maybe even good literature, despite not having been produced as a communication by anyone for anyone."). Arguably, my claim here is too strong. It may not be the case that a string of marks produced by an automaton can (without more) have a sentence meaning; at a minimum, there is the problem of deciding what language those marks are supposed to be in, given that the automaton had no intentions on that score. See Alexander, *supra* note 86, at 360-61 & n.11. However, I take it to be unproblematic that a statute can have a sentence meaning, even though legislators share no mental states beyond their intention to enact into law the marks constitutive of the text of the statute, plus a shared understanding of which set of linguistic conventions (e.g., English) those marks are to be subsumed under. See Waldron, *supra* note 86, at 334-40, 352-56 (arguing for a minimal view of the intentionality necessary to make legislative enactments meaningful). As Waldron puts it: "The intentional speech-acts of the legislature are constitutional functions of the intentional voting-acts of the individual members; but what matters here is simply the intentionality of 'yea' or 'nay' in relation to a given text, not any hopes, aspirations, or understandings that may have accompanied the vote." *Id.* at 353. The necessary conditions for other types of legal utterances (such as judicial orders) to possess sentence meaning are similar, and similarly minimal.

meanings. If the city council enacts an ordinance proclaiming that "Christianity is the religion of our forefathers," then this ordinance at a minimum has the (descriptive) sentence meaning, "Christianity is the religion of our forefathers." If the legislature enacts a statute that provides, "Blacks may not marry outside their race," then this statute at a minimum has the (prescriptive) sentence meaning, "Blacks may not marry outside their race."⁹³

But what, more precisely, does sentence meaning consist in? There are various, conflicting answers to that question current within the philosophy of language.⁹⁴ The view of sentence meaning I will rely upon in this Article is the following: "An utterance x . . . [sentence]-means that p in a community if and only if there prevails in that community a convention to use x in order to [speaker's]-mean that p ."⁹⁵ More crisply, the sentence meaning of a linguistic utterance is what the utterance *conventionally communicates*. Language is understood as a series of conventions for performing speech-acts (that is, actions with speaker's meanings); in turn, the sentence meaning of an action is simply the meaning assigned to it by those conventions.⁹⁶ (A convention is, roughly, an arbitrary regularity such that we each have reason to conform to it, if everyone else does.)⁹⁷ For short, I will call this view the conventionalist view of language. It affirms that speakers do have meanings, indeed that speaker's meaning lies at the foundation of language. Nonetheless, the conventionalist view also recognizes that speakers dramatically benefit by the existence of conventions for performing speech-acts. I do in fact perform a linguistic utterance when I draw a man with a big boot so as to get you to believe that the postman is coming, but it would be much easier for both of us if I could

⁹³ "Prescriptive" meaning is the linguistic meaning that inheres in the creation of legal rights, duties, etc., and is discussed in detail below in Part I.C.

⁹⁴ See BLACKBURN, *SPREADING THE WORD*, *supra* note 55, at 127-34 (describing competing accounts of sentence meaning).

⁹⁵ Avramides, *supra* note 79, at 80 (variables italicized).

⁹⁶ See also BLACKBURN, *SPREADING THE WORD*, *supra* note 55, at 130-34 (presenting a conventionalist account of language); MARTIN, *supra* note 79, at 77-82 (same); Alston, *supra* note 82 (same). These authors would dissent, in various ways, from the particular version of conventionalism I have just presented and upon which I will rely in this Article: An utterance x sentence-means that p only if it is conventional to use x in order to speaker's-mean that p . I rely upon that version for expository purposes, because of its simplicity. Nothing I say hinges upon that version being correct or indeed—as I mention below—on a conventionalist, as opposed to truth-conditional, account of sentence meaning being correct.

⁹⁷ The seminal modern analysis of the concept of convention is DAVID LEWIS, *CONVENTION: A PHILOSOPHICAL STUDY* (1969).

simply say, "The postman is coming."⁹⁸

The conventionalist view is a standard view within the philosophy of language. While it may not be the dominant view, and in any event, is certainly not a consensus view (no view is!), I rely upon it here because it is the view most hospitable to expressive theories of law. If expressive theories fail on the conventionalist view (as I will argue they do), then *a fortiori* they will fail on a view of sentence meaning that is less accommodating to the claim that law is linguistically meaningful. The conventionalist view is particularly accommodating to that claim for the following reason: it can accommodate multiple types of sentence meanings. Speech-acts fall into a variety of categories.⁹⁹ *Descriptive* speech-acts—actions where the speaker intends to get the hearer to believe some proposition about the world—are simply one category. *Directive* speech-acts—actions where the speaker intends to get the hearer to do something—comprise a second and quite different category. If I tell you to "close the door," I do not intend to induce a belief in you; rather, I intend that you act to close the door and that you do so by virtue of your recognition of my intention. Beliefs (the hallmark of descriptive speech-acts) do not enter the analysis of directive speech-acts at all, or at least not in the same way that they enter into the analysis of descriptive speech-acts. Other categories of speech-acts include so-called *commissives* (where the speaker commits himself to doing something, e.g., "I promise to pay you \$100"); *expressives* (where the speaker expresses a psychological state, e.g., "I am sorry about your illness"); and *declarations* (where the speech-act itself constitutes the change in the world to which it refers, e.g., "I declare you husband and wife").¹⁰⁰

⁹⁸ See Avramides, *supra* note 79, at 80-81 (presenting the view that linguistic conventions solve coordination problems).

⁹⁹ See generally JOHN R. SEARLE, *A Taxonomy of Illocutionary Acts*, in *EXPRESSION AND MEANING* 1-29 (1979).

¹⁰⁰ This taxonomy is taken from SEARLE, *id.* at 2-29. See William Croft, *Speech Act Classification, Language Typology and Cognition*, in *FOUNDATIONS OF SPEECH ACT THEORY*, *supra* note 82, at 460 ("The most widely accepted speech act classification is that of Searle . . ."). I use the term "descriptive" rather than "assertive" speech-act to refer to Searle's first category so as to highlight the fact that "[t]he simplest test of an assertive [in my terminology, descriptive] is this: [you can] literally characterize it . . . as true or false." SEARLE, *supra* note 99, at 13. I also have changed Searle's definition of descriptive speech-acts to make it more Gricean. One crucial element of Grice's view is its "perlocutionary" cast, namely, that speech-acts of whatever kind involve an intention to achieve some response (e.g., a mental state, an action) by the listener. See Alston, *supra* note 82, at 33-36; Avramides, *supra* note 79, at 71-78. Searle departs from this feature of Grice's view in his own analysis of descriptives. As he puts it, "[t]he point or purpose of the members of the assertive class is to commit the

The conventionalist view of sentence meaning allows that there can be as many kinds of sentence meanings as there are kinds of speaker's meanings.¹⁰¹ There might well be conventions for uttering directives, conventions for uttering commissives, conventions for uttering expressives, and conventions for uttering declarations, as well as conventions for uttering descriptions. By contrast, competing views of sentence meaning, such as the Davidsonian view that the meaning of a sentence consists (very roughly) in its truth conditions,¹⁰² make good sense of descriptive sentence meaning but have trouble explicating the other variants. Since at least some expressive theories posit that law has nondescriptive meaning—for example, the theory that punishment means condemnation, with “condemning” best understood as some kind of nondescriptive speech-act¹⁰³—a purely descriptivist construal of sentence meaning will prove to be a less secure basis than speech-act conventionalism for expressive theories of law.

C. *Prescriptive Meaning and Further Meaning*

Law surely does possess meaning—sentence meaning—in at least one way. I will call this “prescriptive meaning.” That law is meaningful in at least this way should not be overlooked by critics of expressive theory. Grandiose claims by expressivists about law's symbolism ought not tempt the critic to the opposite extreme—to the conclusion that legal decisions are meaningless.¹⁰⁴ On the other hand, expressivists do not typically confine themselves to prescriptive meaning; they typically claim that legal decisions have *further meaning*, beyond what these decisions prescribe; and this claim about further meaning is, indeed, a fair target for the critic's skepticism.

Legal officials, by their actions, create legal rights, duties, liberties, powers, liabilities, and immunities. They do so via the utterance of

speaker . . . to something's being the case, to the truth of the expressed proposition.” SEARLE, *supra* note 99, at 12. For expository purposes, I stick to a Gricean understanding of speaker's meaning and matched sentence-meaning conventions.

¹⁰¹ See, e.g., BLACKBURN, SPREADING THE WORD, *supra* note 55, at 133 (noting possibility of linguistic conventions for performing nondescriptive speech-acts).

¹⁰² See Donald Davidson, *Truth and Meaning*, 17 SYNTHESE 304 (1967). For a sophisticated summary of Davidson's view, see David Wiggins, *Meaning and Truth Conditions: From Frege's Grand Design to Davidson's*, in A COMPANION TO THE PHILOSOPHY OF LANGUAGE, *supra* note 79, at 3.

¹⁰³ On Searle's taxonomy, condemning is probably an “expressive” speech-act.

¹⁰⁴ Cf. Michael W. McConnell, *Religious Freedom at a Crossroads*, 59 U. CHI. L. REV. 115, 150-51 (1992) (suggesting, in the Establishment Clause context, that claims about the symbolic content of a challenged state action are parasitic upon claims about the justifiability of the action with respect to nonexpressive criteria).

certain sentences: "No person may drive more than fifty miles per hour"; "A will is valid if attested by two witnesses"; "No state shall abridge the right of free speech"; "A probationary employee may be fired at will"; and so on.¹⁰⁵ By prescriptive meaning, I mean just the kind (or kinds) of sentence meaning(s) that these sentences possess. For a legal official or body to utter, "No person may drive more than fifty miles per hour" is to invoke the conventions by which persons are placed under a legal duty not to drive more than fifty miles per hour; for a legal official or body to utter, "A probationary employee may be fired at will" is to invoke the conventions by which employers are accorded a legal liberty to fire probationers.

I have not identified *which* speech-act convention(s) the utterance of a legal sentence invokes and will not try to do so in this Article. "Prescriptive meaning" will not be analyzed; I simply use that term to refer to *whatever* sentence meaning(s) legal utterances have, in virtue of their functioning to create legal rights, duties, liberties, and so on. For one cannot analyze prescriptive meaning, in any detail, without a well worked-out analysis of the concept of "law"—a concept that remains a subject of considerable and continuing debate among legal philosophers.¹⁰⁶ Someone who thinks that law is, paradigmatically, a *command* will be inclined to analyze prescriptive meaning as *directive* meaning.¹⁰⁷ As I explained above, a directive is a speech-act whereby the speaker intends to get the hearer to do something;¹⁰⁸ directive sentences, such as "Give me your money," are conventional mechanisms

¹⁰⁵ See, e.g., CARL WELLMAN, *REAL RIGHTS* 12-37 (1995) (specifying different types of legal positions, such as liberties, claims, powers, and immunities).

¹⁰⁶ The major contributions to the modern, jurisprudential literature about the concept of law include, of course, H.L.A. HART, *THE CONCEPT OF LAW* (1961) [hereinafter HART, *THE CONCEPT OF LAW*], and also, among others, JOSEPH RAZ, *THE CONCEPT OF A LEGAL SYSTEM* (1970); JOHN FINNIS, *NATURAL LAW AND NATURAL RIGHTS* (1980); and RONALD DWORKIN, *LAW'S EMPIRE* (1986). The modern debate focuses, in particular, on the difference between "positivist" and "natural law" accounts of the concept of law and on the truth of the two types of accounts. Recent contributions to this debate include H.L.A. HART, *THE CONCEPT OF LAW* (2d ed. 1994); *Special Issue: Postscript to H.L.A. Hart's The Concept of Law, Part 1*, 4 *LEGAL THEORY* 249 (1998); *Special Issue: Postscript to H.L.A. Hart's The Concept of Law, Part 2*, 4 *LEGAL THEORY* 381 (1998); *THE AUTONOMY OF LAW: ESSAYS ON LEGAL POSITIVISM* (Robert P. George ed., 1996); and W.J. WALUCHOW, *INCLUSIVE LEGAL POSITIVISM* (1994). A good introduction is BRIAN BIX, *JURISPRUDENCE: THEORY AND CONTEXT* (1996). See also Adler, *supra* note 47, at 81-87 (discussing and citing contributions to a related, ongoing debate about the nature of authority).

¹⁰⁷ See HART, *THE CONCEPT OF LAW*, *supra* note 106, at 18-25 (presenting a model based on Austin's view of law as the sovereign's coercive orders).

¹⁰⁸ See *supra* text accompanying notes 99-100 (summarizing conditions for descriptive, directive, commissive, expressive, and declarative speech-acts).

for performing directive speech-acts; and the simple-minded command theorist will thus propose that, when a legal body such as a legislature, court, or agency utters a duty-creating sentence ("No person may drive more than fifty miles per hour"), this legal action simply invokes the ordinary, directive conventions prevalent within the linguistic community. The notorious problem with this analysis is that it accords precisely the same type of meaning to "No person may drive more than fifty miles per hour" as it accords to "Give me your money"; it draws no distinction between the utterances of a legal official and the utterances of a robber.¹⁰⁹ Further, it seemingly fails to explain the meaning of certain other legal sentences, such as those that create liberties. "Landowners are at liberty to exclude trespassers by means of fences, warnings, threats, and other nonviolent means" does not, apparently, direct landowners to do anything.¹¹⁰

A partial response to these problems, at least to the first, might be to append a *descriptive* sentence to legal directives. On this refinement of the command theory, "No person may X"—when uttered by a legal official—has the hybrid meaning: "Do not X" (a command) and "It is morally required that persons not X" (a description). As Joseph Raz has argued, what distinguishes the legal official from the robber is that the legal official claims moral authority for her commands.¹¹¹ More radically, in line with a natural-law analysis of the concept of law, one might propose that legal sentences *merely* describe moral facts. On such a view, the utterance of "No person may X" by a legal official means "Persons are under a moral obligation not to X"; the utterance of "Persons are free to X" means "Persons have no moral obligation not to X"; and so on. (Note how this solves, in a direct way, the problem of accounting for the meaning of sentences that create legal liberties.) In fact, Heidi Hurd has advanced a natural law theory of prescriptive meaning along purely descriptivist lines.¹¹²

So prescriptive meaning is difficult to analyze, but legal sentences

¹⁰⁹ See, e.g., JOSEPH RAZ, *THE MORALITY OF FREEDOM* 26 (1986) ("[A]ll political authorities must and do resort to extensive use of and reliance on coercive and other threats. Yet it is clear that all legal authorities do much more. They claim to impose duties and to confer rights.").

¹¹⁰ See, e.g., HART, *THE CONCEPT OF LAW*, *supra* note 106, at 77-79 (summarizing objections to the view of law as the sovereign's coercive orders); WELLMAN, *supra* note 105, at 17-24 (discussing the possibility of liberty-creating and power-creating laws that are not best understood as fragments of duty-creating laws).

¹¹¹ See RAZ, *supra* note 109, at 26. But see Matthew H. Kramer, *Requirements, Reasons, and Raz: Legal Positivism and Legal Duties*, 109 *ETHICS* 375 (1999) (analyzing and criticizing Raz's view).

¹¹² See Hurd, *supra* note 72, at 990-1028.

surely do possess it, whatever precisely it is. They surely do invoke certain linguistic conventions in creating legal rights, duties, etc.—whether those conventions are best understood as directive, descriptive, directive and descriptive, or something else. Again: the view that law is meaningless is too extreme. But the existence of a characteristically legal meaning does not license the further claim that legal sentences possess whatever *nonprescriptive meaning* is posited by the expressive theorist. Expressive theories are typically dualist; they typically presume not only that legal officials create rights, duties, liberties, etc., but also that these officials stigmatize, endorse, condemn, express respect or contempt, or utter some other kind of sentence meaning that is not entailed by the sheer act of prescribing.¹¹³ “No blacks may own automobiles” is thought *both* to direct blacks not to possess automobiles (on a simple command theory of law) *and* to describe or declare blacks as social inferiors.¹¹⁴ “The defendant is committed to the state penitentiary” is thought *both* to direct the defendant to the penitentiary (on a command theory) *and* to condemn his wrongdoing.¹¹⁵

To be sure, legal officials have the capacity to utter nonprescriptive sentences.¹¹⁶ For example, the city council could simply issue a proclamation that “Christianity is the official religion” (a proclamation that does not confer legal rights, duties, liberties, etc. on anyone);¹¹⁷ it is not constrained to endorse by prescribing. Similarly, a judge could accompany his sentence of conviction with an additional

¹¹³ See *infra* text accompanying notes 162, 223, 249-55, 300-01 (describing nonprescriptive meanings typically posited by expressivists about punishment, race discrimination, religion, and regulation).

¹¹⁴ See, e.g., KOPPELMAN, *supra* note 32, at 57-76 (summarizing stigma theory).

¹¹⁵ For a particularly clear presentation of the view that the hard treatment partly constitutive of (normal) punishment has a further, condemnatory content, see DUFF, *supra* note 28, at 235-45. I consider, below, the possibility of a purely expressive punishment—what I call *Punishment_E*. See *infra* text accompanying note 164. *Punishment_E* would lack prescriptive meaning, and thus does not falsify my claim here that expressivists about punishment posit a nonprescriptive meaning attached to the utterances of legal officials.

¹¹⁶ See Lewis A. Kornhauser, *No Best Answer?*, 146 U. PA. L. REV. 1599, 1626-34 (1998) (noting that legal officials have expressive resources other than the announcement or implementation of legal rules, but suggesting that advocates of an expressive theory of law are concerned primarily with the expressive content of rule announcement and implementation).

¹¹⁷ See Steven B. Epstein, *Rethinking the Constitutionality of Ceremonial Deism*, 96 COLUM. L. REV. 2083, 2095, 2092 (1996) (analyzing Establishment Clause status of “ceremonial deism,” defined *inter alia* as “practices involving . . . prayer, invocation, benediction, supplication, appeal, reverent reference to, or embrace of, a general or particular deity [that are] created, delivered, sponsored, or encouraged by government officials,” such as the designation of “In God We Trust” as our national motto).

sentence, contained in the text of his opinion, that condemns the defendant's wrongdoing.¹¹⁸ Furthermore, it is possible that certain straight prescriptive sentences *do* possess dual meanings; "No blacks may own automobiles" might really invoke a convention both for prescribing and for stigmatizing. In either case, however, the attachment of nonprescriptive meaning to the utterances of legal officials is an empirical and contingent matter, not entailed by the fact of law. This point will prove of some significance in Part II, when we turn to evaluate expressive theories.

At this juncture, I should mention the possibility of nonliteral sentence meaning—a possibility that is helpful to the expressivist.¹¹⁹ An action *A* has sentence meaning *M* on the definition I am using in this Article if—given existing conventions in the relevant society—this action is a conventional mechanism for performing a speech-act with speaker's meaning *M*. Some theorists of language have suggested that an action can have a conventional but nonliteral ("pragmatic" rather than "semantic") linguistic content. For example, the action of uttering "Do you think you could turn down the radio?" has the literal, conventional meaning of a question, but also possesses an arguably nonliteral and arguably conventional content to the effect, "Turn down the radio."¹²⁰ I am not sure whether the distinction between literal and nonliteral conventional meaning is coherent; if *M*₁ and *M*₂ are both conventional meanings of action *A*, then why are not those both *A*'s literal meaning?¹²¹ In any event, the point to emphasize here is that *A*'s sentence meaning, by my definition, encompasses any kind of linguistic meaning (literal or not) for the communication of which *A* is conventional. This broad definition helps the expressivist, be-

¹¹⁸ Cf. *infra* text accompanying notes 432-39 (stating that legal utterances, or some class thereof, should arguably be accompanied by a public justification).

¹¹⁹ I am indebted to Larry Alexander and Michael Moore for suggesting to me that the kind of linguistic meaning typically described by expressivists—e.g., the stigmatic meaning of discrimination—could be a nonliteral or pragmatic, but conventional, meaning; and to Leo Katz for generally urging me to incorporate the possibility of nonliteral meaning into my account.

¹²⁰ See PAUL GRICE, *Logic and Conversation*, in *STUDIES IN THE WAY OF WORDS*, *supra* note 1, at 37-40 (discussing possibility of "generalized conversational implicature"); SEARLE, *Indirect Speech Acts*, *supra* note 82, at 49 ("[T]here can be conventions of usage that are not meaning conventions. I am suggesting that 'can you', 'could you', 'I want you to', and numerous other forms are conventional ways of making requests . . . but at the same time they do not have an imperative [literal] meaning . . ."). For a general discussion of nonliteral meaning, see KENT BACH & ROBERT M. HARNISH, *LINGUISTIC COMMUNICATION AND SPEECH ACTS* 155-72 (1979).

¹²¹ See MARTIN, *supra* note 79, at 218-21 (arguing against the view that words, as opposed to speakers, can have a metaphorical meaning).

cause it supports the view that a legal utterance can possess both a prescriptive sentence meaning and the nonprescriptive sentence meaning targeted by the expressivist. It supports the view that "No blacks may own automobiles" could sentence-mean both "Blacks are hereby directed not to possess automobiles" and "Blacks are inferior." Whether prescriptive utterances have a further, nonprescriptive content is an empirical and contingent matter, both in the sense that this content is not entailed by the act of prescribing, and in the sense that nothing in my definition of sentence meaning precludes attaching a further (and perhaps "nonliteral") sentence meaning to the act of prescribing.

D. *What Makes a Theory Genuinely Expressive?*

We now turn to a conceptual problem of great importance for the debate about expressive theories of law, one that the existing literature has failed to clarify: What makes a moral theory genuinely expressive? On virtually any moral theory, the linguistic meaning (specifically, the sentence meaning) of a legal official's action can have a moral impact. What distinguishes this trivial and virtually universal feature of moral theories from the nontrivial characterization of a theory as "expressive"?

To see the issue here, imagine that my moral theory tells me to minimize the aggregate number of premature deaths.¹²² (It is a crude kind of consequentialist theory.¹²³) Imagine further that if governmental officials utter a certain sentence S_0 at a particular time—where S_0 is, let us assume, the sentence, "Premature death is a bad thing"—this utterance will cause the aggregate number of premature deaths to be minimized. Other utterances at that time, or the performance of linguistically meaningless actions, will lead to a greater number of deaths. The details of the causal story do not matter; perhaps killers

¹²² Cf. LESTER B. LAVE, *THE STRATEGY OF SOCIAL REGULATION: DECISION FRAMEWORKS FOR POLICY* 15-17 (1981) (describing and discussing "risk-risk" decision procedure, whereby regulatory agencies seek to minimize the aggregate number of deaths); Cass Sunstein, *Health-Health Tradeoffs*, 63 U. CHI. L. REV. 1533 (1996) (same); W. Kip Viscusi, *Economic Foundations of the Current Regulatory Reform Efforts*, 10 J. ECON. PERSP. 119, 129-31 (1996) (same).

¹²³ See generally CONSEQUENTIALISM AND ITS CRITICS (Samuel Scheffler ed., 1988) (analyzing the structure of consequentialist moral theories); SHELLY KAGAN, *THE LIMITS OF MORALITY* (1989) [hereinafter KAGAN, *THE LIMITS OF MORALITY*] (same); SHELLY KAGAN, *NORMATIVE ETHICS* (1998) [hereinafter KAGAN, *NORMATIVE ETHICS*] (same); SAMUEL SCHEFFLER, *THE REJECTION OF CONSEQUENTIALISM* (rev. ed. 1994) (same).

would be persuaded by the utterance of S_0 to refrain from killing, perhaps crime victims would be inspired to take additional precautions, perhaps doctors would redouble their care.¹²⁴ In any event, the official utterance of S_0 is morally required by our crude consequentialist theory; what law means, then, has a moral impact. But surely we have not yet discovered a genuine "expressive theory of law." Governmental officials are simply obliged to do whatever it takes to minimize premature death and are prohibited from doing anything else. The fact that uttering S_0 leads to minimization is purely contingent and causal. If this crude consequentialist theory counts as "expressive" just because certain meanings happen to have the right (or wrong) sort of consequences, then *any* nonrigged moral theory of governmental action will count as "expressive" and the description will be trivial.¹²⁵

The point has both theoretical and practical importance. It is theoretically important, because expressivists (or at least some of them) claim to be offering a new kind of moral theory and specifically to be criticizing consequentialism.¹²⁶ If the description "expressive"

¹²⁴ See Kahan, *supra* note 9, at 602-05 (describing various mechanisms whereby governmental expression can cause reductions in future crime).

¹²⁵ This point is, in effect, articulated by HART, *supra* note 27, who distinguishes between expressivism as a distinctive type of theory of punishment, and expressivism as a claim about the role of denunciation within a nonexpressive theory. Hart writes:

Notwithstanding the eminence of its legal advocates, [the expressive] justification of punishment, especially when applied to conduct not harmful to others, seems to rest on a strange amalgam of ideas. It represents as a value to be pursued at the cost of human suffering the bare expression of moral condemnation But is this really intelligible? . . .

It is, I think, probable that what the advocates of this theory really mean by . . . the 'appropriate' expression of moral condemnation is one that is effective in instilling or strengthening in the offender and in others respect for the moral code which has been violated. But then the theory assumes a different character; it is no longer the theory that the legal enforcement of morality is a value apart from its consequences; it becomes the theory that the legal enforcement of morality is valuable because it preserves an existing morality.

Id. at 65-66.

¹²⁶ See, e.g., DUFF, *supra* note 28, at 151-86, 233-66 (explicitly criticizing consequentialism and advocating a nonconsequentialist, expressive theory); NOZICK, *supra* note 28, at 363-97 (same); Pildes & Anderson, *Slingshot Arrows*, *supra* note 8 (same); Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 66-72 (same); Primoratz, *supra* note 28, at 196-98 (same). The differentiation between expressivism and nonexpressive theories (paradigmatically consequentialism) is less explicit in the constitutional law scholarship discussed *infra* Part II.B, but the theories of race discrimination and the Establishment Clause defended by constitutional expressivists are in fact genuinely expressive by the criteria articulated in this Section.

applies to any nonrigged theory, the claim and the critique are misguided. As for practical importance: the statements required or prohibited by, say, ordinary consequentialism would presumably be quite different from those required or prohibited by a genuine expressive theory.¹²⁷ (For example, ordinary consequentialists would find it difficult to explain why government should always respond to moral wrongdoing by condemning it.¹²⁸) Anyone interested in deciding what government ought, or ought not, to say should determine whether there exist genuine expressive constraints and goals, or simply the ordinary constraints and goals of a nonexpressive theory that happen, in a particular instance, to prohibit or require expression.

How, then, should a genuinely expressive theory of law be characterized? One temptation, to be avoided, is to say that a theory is genuinely expressive if and only if the linguistic meaning of an action can have *noncausal* moral impact. This would rule out the crude consequentialist theory just described. But the causal/noncausal distinction would not properly handle the case where (1) the meaning of a particular action has a noncausal rather than a causal impact; and yet (2) linguistic meaning plays no deeper role in the theory than in the case of a mere causal impact. Imagine that our theory is preference-utilitarianism, which enjoins government to maximize the satisfaction of preferences.¹²⁹ Imagine further that, as it happens, some citizen prefers that government says S_0 . In this scenario, the meaning of governmental utterances has a noncausal moral impact: government's utterance of S_0 is morally required, *ceteris paribus*, not because of what it causes, but because of what it constitutes, namely the satisfaction of someone's preference. Yet the possibility of such a scenario hardly makes preference-utilitarianism a genuine expressive theory. Indeed, expressivists typically have taken preference-utilitarianism as their chief target.¹³⁰

I propose, instead, that the concept of an "expressive theory" be defined as follows:

¹²⁷ See HART, *supra* note 27, at 67 (noting that, if an expressive response to wrongdoing is claimed to be morally justified as instrumental to the reinforcement of existing morality, that claim is open to criticism on factual grounds: "the assertion that [expression] does operate in the manner supposed . . . requires evidence in support, and at least in relation to sexual morality there is little to be found").

¹²⁸ See *infra* text accompanying notes 442-43.

¹²⁹ See *infra* text accompanying notes 337-40 (explicating preference-utilitarianism in detail).

¹³⁰ See, e.g., Pildes & Anderson, *Slingshot Arrows*, *supra* note 8; Sunstein, *Incommensurability and Valuation*, *supra* note 16.

"Expressive Theory of Law": A Definition

An "expressive moral theory" is a theory such that the moral factors $F_1 \dots F_n$ that collectively determine the moral status of an action (the status of the action as required, prohibited, or permitted)¹³¹ include at least one expressive factor. A moral factor F_i is "expressive" if it refers to the linguistic meaning of an action. More precisely, a moral factor F_i is expressive if *meaningless* actions (actions lacking linguistic meaning) must fare equally well, with respect to the factor. An "expressive theory of law" is a moral theory such that the moral factors $F_1 \dots F_n$ that collectively determine the moral status of a *legal official's* action include at least one expressive factor.

Any moral theory will specify certain moral factors that combine to determine the status of a given action as required, prohibited, or permissible, under that theory. By a "moral factor," I mean a way of characterizing an action that has a canonical or foundational status within the theory—one that figures in the best statement of the theory¹³² and cannot be reduced to other factors.¹³³ For example, utilitarianism specifies a single moral factor F_w : the amount of well-being produced by a given action.¹³⁴ A more sophisticated moral theory, which recognizes the existence of constraints upon welfare-maximizing action, will specify multiple moral factors: F_w plus $F_{D1}, F_{D2}, \dots, F_{Dn}$, where F_{Di} might

¹³¹ If there are additional moral statuses beyond these three that an action can possess, e.g., the status of supererogatory, see generally DAVID HEYD, *SUPEREROGATION: ITS STATUS IN ETHICAL THEORY* (1982), the definition here can readily be amended to reflect them.

¹³² I will not attempt to present specific criteria for "best," "canonical," or "foundational," for that goes deeper than is necessary into the nature of moral theorizing; whatever the criteria, an expressive factor is one which figures in the best statement of the theory, thus specified.

¹³³ See generally KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 25-186 (presenting a template for moral theories that incorporate "moral factors" in this sense). Cf. Russ Shafer-Landau, *Moral Rules*, 107 *ETHICS* 584 (1997) (discussing the possibility of "particularistic" theories that do not incorporate such "moral factors"). Although my definition of genuine expressivism could be modified to accommodate the possibility of a genuinely expressive, particularistic theory—as a theory such that, on certain particular occasions, the (nonuniversal) properties identified as morally relevant will be expressive properties—I think that particularism is problematic. I will therefore focus on explicating expressivism within a nonparticularistic framework. Further, as far as I can tell, the specific expressivists whose theories I present and criticize in this Article are not particularists.

¹³⁴ For a discussion of utilitarianism, see GEOFFREY SCARRE, *UTILITARIANISM* (1996); *UTILITARIANISM AND BEYOND* (Amartya Sen & Bernard Williams eds., 1982); and J.J.C. SMART & BERNARD WILLIAMS, *UTILITARIANISM: FOR AND AGAINST* (1973).

be a constraint against intentionally causing serious harm,¹³⁵ F_{D2} might be a constraint against breaching promises,¹³⁶ and F_{D3} might be a constraint against lying.¹³⁷

F_w is a nonexpressive factor: it concerns the total amount of well-being produced by an action, not the action's linguistic meaning. (Two meaningless actions can produce different amounts of overall well-being.) Similarly, the factor F_{D1} is nonexpressive: what matters with respect to this factor is whether an action causes serious harm and whether the mental state behind the action is intentional, not whether the action has a particular meaning. (Given two meaningless actions, one can constitute a serious, intentional harming, while the other does not.) The same story can be told about factor F_{D2} , the constraint against promise-breaching. By contrast, the factor F_{D3} in this mini-theory is expressive. The constraint against lying refers to the linguistic meaning of actions. It delineates a property of action (the property of constituting a lie) such that only a linguistically meaningful action can possess that property. Equivalently, two meaningless actions will necessarily fare the same with respect to the factor F_{D3} . Given two meaningless actions, neither will trigger the no-lying constraint, because it is a necessary condition for an action to constitute a lie that it be linguistically meaningful.¹³⁸

To recapitulate: Moral theories posit one or more basic factors that collectively determine the moral status of actions (here, specifically, the actions of legal officials). An expressive theory of law is a theory such that, among the basic factors governing the actions of legal officials, at least one such basic factor is "expressive" in the sense that linguistically meaningless actions must fare identically with respect to that factor.

My definition of an expressive theory accomplishes the task of making the description "expressive" an interesting and nontrivial one. It distinguishes between genuine expressivism and the mere ascription of moral impact to linguistic meaning—an ascription which, again, virtually any moral theory can warrant. Linguistic meaning has a "moral impact," in a particular choice situation, if linguistic meaning *co-varies* with some moral factor F_i —if, given two actions that differ with respect to F_i , the actions also differ with respect to linguistic

¹³⁵ See KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 70-105.

¹³⁶ See *id.* at 116-25.

¹³⁷ See *id.* at 106-16.

¹³⁸ See *infra* text accompanying notes 393-415 (further discussing constraints against lying and harming).

meaning. To return to the example at the beginning of this section: imagine that F_i is the total number of deaths produced by an action and that, in a particular choice situation, an action with the linguistic meaning "Premature death is a bad thing" produces fewer deaths than an action with no linguistic meaning. Linguistic meaning has a moral impact in this choice situation. It *matters* in light of F_i whether the agent utters "Premature death is a bad thing" as opposed to saying nothing. But the factor F_i is not an expressive factor, since two meaningless actions can produce different numbers of overall deaths. The performance of the utterance is here morally required, not by virtue of what the utterance means—not by virtue of the sentence-meaning conventions that it triggers—but by virtue of its effect on the overall number of deaths.

The definition I offer here not only makes the description, "expressive," nontrivial and interesting, but also tracks our pretheoretical sense of what an expressive theory is. Intuitively, theories such as the death-minimization theory, or preference-utilitarianism, or a theory that prohibits the intentional causing of serious harm are not genuinely expressive because a characterization of a particular action in terms of its linguistic meaning can always be reduced to a characterization that does not refer to linguistic meaning. Linguistic meaning lacks *basic* moral relevance within such theories; the most perspicuous description of an action will refer to its effect on overall deaths, or on preference-satisfaction, or to the serious harm it causes, not to what it means. My definition of an "expressive theory" is meant to serve as a somewhat fuller and more precise articulation of this pretheoretical intuition.

Finally, I should note that the concept of an "expressive factor," upon which my definition crucially depends, can be fleshed out in different ways. Roughly speaking, a factor F_i is "expressive" if it refers to the linguistic meaning of actions. If F_1 is a factor that looks to whether government's action constitutes a condemnation, and F_2 is a factor that looks to whether government's action constitutes a lie, while F_3 is a factor that looks to the overall degree of preference-satisfaction produced by government's actions, then F_1 and F_2 but not F_3 would seem to be expressive. But what, precisely, is the feature of F_1 and F_2 that makes them, and not F_3 , "refer" to linguistic meaning?¹⁹⁹ Shall we say (1) that linguistically identical actions must fare identically with re-

¹⁹⁹ Cf. BLACKBURN, *SPREADING THE WORD*, *supra* note 55, at 302-45 (summarizing debates about how language refers to individual objects).

spect to F_1 and F_2 ; or, more weakly, (2) that meaningless actions must fare identically with respect to F_1 and F_2 ? I have chosen the second and weaker option: a factor F_i is expressive, on my account, if two meaningless actions must fare identically with respect to the factor. The problem with the first and stronger option is that it actually excludes F_2 : given two actions with identical linguistic content, one may be a lie and the other not, depending on facts about the world. For example, "grass is green" is true in a world where grass is green, but false in a world where grass is blue, even though "grass is green" means the same thing in both worlds.¹⁴⁰

The formulation of "expressive factor," offered here, seeks once more to track our pretheoretical intuitions; intuitively, F_2 is expressive. It also serves to make the critical assessment of expressive theories provided below in Parts II and III particularly robust. If the larger set of moral theories that meet my definition of "expressive theory" are unpersuasive, then *a fortiori* the smaller and included set of moral theories that meet a more demanding definition will be unpersuasive.

Cass Sunstein, in his article, *On the Expressive Function of Law*, confronts the problem I have addressed in this section. His answer is somewhat different from mine. Sunstein distinguishes not between

¹⁴⁰ Certainly this is true on a traditional "correspondence" theory of truth. See RICHARD L. KIRKHAM, *THEORIES OF TRUTH: A CRITICAL INTRODUCTION* 119-40 (1992) (describing variants of correspondence theory).

It might be objected that moral factors, such as expressive factors, are relevant for choosing between different actions in the same choice situation, and that the putative problem I have just identified, concerning the strong definition of "expressive factor," would not in fact arise as between different actions in the same choice situation. That is: requiring that linguistically identical actions in the same choice situation fare identically with respect to an "expressive factor" does track our intuitive sense of what an "expressive factor" involves, and does count F_2 (the no-lying factor) as expressive. My response to this objection is as follows: I agree that moral factors are relevant for choosing between different actions in the same choice situation; my own (weak) definition of "expressive factor" should be understood to require only that meaningless actions in the same choice situation fare identically with respect to such a factor, not that two different meaningless actions in two different choice situations do so. But the stronger definition of "expressive factor" does, seemingly, prove problematic even with respect to linguistically identical actions in the very same choice situation. Imagine choosing between A_1 , which means "the grass will be red at noon tomorrow," and A_2 , which means "the grass will be red at noon tomorrow," where A_1 but not A_2 will produce a world in which the grass will be red at noon tomorrow. Then A_2 but not A_1 is a lie—since A_2 but not A_1 is a false statement of fact—even though A_2 and A_1 have the very same linguistic meaning.

These considerations led me to employ a weaker definition of "expressive factor." As I say below, if expressive theories in this weaker sense are unpersuasive, then *a fortiori* expressive theories in the stronger sense are. So I do not see much to lose by weakening the definition.

theories that incorporate expressive factors and theories that merely warrant the ascription of moral impact to linguistic meaning, but rather between (1) theories where the meaning of law has "intrinsic" value, and (2) theories where the meaning of law merely has a "consequential" effect:

I have suggested that some expressivists are concerned with norm management, whereas others are concerned with the "statement" law makes entirely apart from its consequences. . . . For those who endorse the expressive function of law, the most important testing cases arise when (a) people support laws because of the statement made by such laws but (b) the effects of such laws seem bad or ambiguous, even by reference to the values held by their supporters My basic proposition is that, at least for purposes of law, any support for "statements" should be rooted not simply in the intrinsic value of the statement, but also in plausible judgments about its effect on social norms and . . . its consequences [more generally].¹⁴¹

If by "intrinsic" Sunstein means noncausal, and by "consequential" he means causal, then his distinction is just the causal/noncausal distinction I criticized above. If, alternatively, by "consequential" Sunstein means to refer to the kind of theory that I have called consequentialist—roughly, a theory similar in structure to utilitarianism—then Sunstein's distinction is problematic for a different reason. A moral theory might be nonconsequentialist and yet still not genuinely expressive.¹⁴² That is true, for example, of a moral theory that incorporates a nonconsequentialist but nonexpressive constraint against intentional killings. The meaning of an action can have an impact with respect to that constraint—if I tell someone the wrong thing, I might kill him—and yet this theory is no more genuinely "expressive" than preference-utilitarianism, which also permits linguistic meaning to have a moral impact. So I will adhere to my own definition of an "expressive theory."

E. *Expressive Theories of Governmental Action and Expressive Theories of Individual Action*

Finally, let me draw a distinction between expressive theories of individual action and expressive theories of law, or of governmental ac-

¹⁴¹ Sunstein, *On the Expressive Function of Law*, *supra* note 20, at 2045; see also Kornhauser, *supra* note 116, at 1622-34 (drawing a similar distinction).

¹⁴² There is nothing in the concept of a nonconsequentialist or, equivalently, a "deontological" constraint that entails expressivity. For a recent, precise account of the structure of deontological constraints, see David McNaughton & Piers Rawling, *Agent-Relativity and the Doing-Happening Distinction*, 63 PHIL. STUD. 167 (1991).

tion. It is particularly important to draw this distinction because the philosopher Elizabeth Anderson has recently advanced the first kind of theory, primarily in her book, *Value in Ethics and Economics*.¹⁴³ Anderson's book has been widely cited and discussed, not only in the philosophical literature on choice and rationality,¹⁴⁴ but also in the legal literature on expressive theories of law.¹⁴⁵ Indeed, Anderson has coauthored a law review article with Professor Pildes, one of the leading legal expressivists,¹⁴⁶ and recently participated in a symposium on *Social Norms, Social Meaning, and the Economic Analysis of Law*.¹⁴⁷ It thus seems to be the case that Anderson herself holds both an expressive theory of individual action and an expressive theory of law. But in any event, the two kinds of theories are conceptually distinct—the first might be true and the second false, or vice versa—and Anderson's oft-cited book is largely a defense of an individual-level theory.¹⁴⁸

Consider the following précis of the theory presented in her book:

The theory of rational action that I propose . . . can be called an expressive theory. An expressive theory defines rational action as action that adequately expresses our rational attitudes toward people and other intrinsically valuable things. According to the rational attitude theory of value, something is valuable if and only if it is rational for someone to value it, to assume a favorable attitude toward it. And to adequately care about something requires that one express one's valuations in the world, to embody them in some social reality.¹⁴⁹

This précis illustrates the ambiguity of the term, "expressive," for the theory just summarized is a different kind of theory than two other kinds, already discussed in this Article, which are also named by that

¹⁴³ ELIZABETH ANDERSON, *VALUE IN ETHICS AND ECONOMICS* (1993) [hereinafter ANDERSON, *VALUE IN ETHICS AND ECONOMICS*]. For a more recent statement of Anderson's views, see Elizabeth Anderson, *Practical Reason and Incommensurable Goods*, in INCOMMENSURABILITY, *supra* note 16, at 90-109.

¹⁴⁴ See Symposium on Elizabeth Anderson's *Value in Ethics and Economics*, 106 ETHICS 508 (1996).

¹⁴⁵ A search of Westlaw's law review database ("JLR") on April 25, 2000, turned up 109 citations to Anderson's book.

¹⁴⁶ See Pildes & Anderson, *Slingshot Arrows*, *supra* note 8.

¹⁴⁷ See Richard A. Posner, *Social Norms, Social Meaning, and Economic Analysis of Law: A Comment*, 27 J. LEGAL STUD. 553, 558-62 (1998) (discussing an unpublished paper by Anderson).

¹⁴⁸ The final chapter of the book, on cost-benefit analysis, does incorporate some of the elements of the expressive theory of regulation that is presented by Pildes & Anderson, *Slingshot Arrows*, *supra* note 8. See ANDERSON, *VALUE IN ETHICS AND ECONOMICS*, *supra* note 143, at 190-220; *infra* Part II.C (discussing this expressive theory of regulation).

¹⁴⁹ ANDERSON, *VALUE IN ETHICS AND ECONOMICS*, *supra* note 143, at 17.

term. First, Anderson's theory is not metaethical expressivism—it is not the kind of theory, concerning the nature of morality, defended by Ayer, Hare, Blackburn, and Gibbard¹⁵⁰—because there is not a hint of noncognitivism in Anderson's book. Anderson does not want to deny the existence of moral or evaluative truth; rather, she argues that attitudes can be truly rational and that it can be truly rational to act so as to express them.¹⁵¹ Second, the theory to which the above *précis* refers is not, and does not entail, an expressive theory of law. To say (a) that individuals should act to express their rational attitudes does not entail (b) that governmental bodies should act to express *government's* rational attitudes.¹⁵²

As I have already explained, it is problematic to characterize multimember governmental bodies as possessing the mental states and, in particular, the attitudes (intentions, purposes, motives, beliefs, preferences) that are unproblematically possessed by individuals.¹⁵³ One might structure Anderson's theory to get around this problem; one might say that, in choosing between A_1 and A_2 , governmental bodies should choose the action with the sentence meaning that best expresses the attitude(s) that it is rational for the members of these bodies to hold. But why is this the appropriate course for governmental bodies to follow? The choice between A_1 and A_2 is a *governmental* choice. It is not, or may not be, the same kind of choice that individuals face (for example, it may have much more dramatic consequences for the well-being of many persons than an individual's choice typically has). Further, individuals have their own lives to lead. To put the point a bit more technically, individuals have their own nonmoral projects, such as their friendships, careers, properties, or families, that they are morally licensed to pursue, within broad moral constraints, notwithstanding overall consequences.¹⁵⁴ (For example, I

¹⁵⁰ See *supra* Part I.A (discussing metaethical expressivism).

¹⁵¹ Among other things, Anderson's book "propose[s] a pragmatic account of how we can objectively justify our value judgments, and . . . defend[s] it against several levels of subjectivist and skeptical criticism." ANDERSON, *VALUE IN ETHICS AND ECONOMICS*, *supra* note 143, at 91.

¹⁵² This point is made by Richard Craswell in his article, *Incommensurability, Welfare Economics, and the Law*, 146 U. PA. L. REV. 1419, 1461 (1998), and by Lewis A. Kornhauser, *supra* note 116, at 1633-34.

¹⁵³ See *supra* text accompanying notes 83-89.

¹⁵⁴ This is now a standard theme within the philosophical literature criticizing utilitarianism. See, e.g., SCHEFFLER, *supra* note 123, at 14-79; BERNARD WILLIAMS, *A Critique of Utilitarianism*, in SMART & WILLIAMS, *supra* note 134, at 108-18; Thomas Nagel, *Autonomy and Deontology*, in CONSEQUENTIALISM AND ITS CRITICS, *supra* note 123, at 142, 145-56; see also KAGAN, *THE LIMITS OF MORALITY*, *supra* note 123, at 3 (noting that or-

am morally licensed to continue my art career, even if it would be better for art and the world if I gave it up and served as the assistant to you, the more artistically talented one.) By contrast, government and government officials acting in an official capacity have no projects of their own in this sense.¹⁵⁵

Imagine that action A_1 will dramatically improve the life-prospects of one group of persons and set back the life-prospects of a second, while action A_2 will dramatically set back the life-prospects of the first group and improve the life prospects of the second. Assume further that individuals are, within broad constraints, licensed to assume an attitude of *loyalty* to other persons or groups, and to choose actions that express loyalty. I can give a loan to my friend, even if the new family that just moved to town could make better use of the money. This hardly shows that government, in choosing between A_1 and A_2 here, should choose the action that has the sentence meaning of expressing an appropriate loyalty. First, the choice is a dramatic one, perhaps sufficiently dramatic that even a private individual making a choice with those consequences would be obliged to ignore her private loyalties and decide impartially.¹⁵⁶ Second, government has no legitimate loyalties, or at least no legitimate loyalties as between different groups of citizens. (Assume that the A_1 - A_2 choice affects only citizens.) Rather, quite plausibly, government should choose between A_1 and A_2 in a way that is fair, or that maximizes overall well-being, or that maximizes the resources of the least-well-off, or that satisfies some other such nonexpressive criteria. It is fully consistent with the fact of individual loyalty, and with the existence of a rational, even moral, duty requiring individuals to express their loyalties, that the criteria

dinary morality confers upon agents "the option of performing (or not performing) acts which from a neutral perspective are less than optimal").

¹⁵⁵ See, e.g., DWORKIN, *supra* note 106 ("We each claim a personal point of view, ambitions and attachments of our own we are at liberty to pursue, free from the claims of others to equal attention, concern, and resource. . . . [But] [w]e allow officials acting in their official capacity no such area at all.").

¹⁵⁶ As Scheffler states:

On a plausible view [of the appropriate modification to consequentialism to accommodate personal projects], the answer to the question of whether an agent was required to promote the best overall outcome in a given situation would depend on the amount of good he could thereby produce . . . and on the size of the sacrifice he would have to make in order to achieve the optimal outcome. More specifically, I believe that a plausible agent-centred prerogative would allow each agent to assign a certain proportionately greater weight to his own interests

SCHEFFLER, *supra* note 123, at 20.

for governmental choice are nonexpressive.¹⁵⁷

Conversely, an expressive theory of law does not entail an expressive theory of individual action. Consider a classic example of the first kind, namely the expressive theory of punishment. One might think that government should condemn wrongdoing, without thinking that individual action is about expression at all. For example, individual lives might be about *being* loyal, *being* a good parent, *being* accomplished, not about expressing those virtues; wrongdoing might be a matter of unjustifiably harming other individual's lives; and punishment might be a collective disapproval of wrongdoing, which governmental officials are required to voice, without it being the case that any private individual is obliged to express her disapproval of the wrongdoing (for each such individual might be morally licensed not to do so, given the demands of her own projects).

I do not deny that there could be links, short of entailment, between expressive theories of individual action and expressive theories of law. The arguments for the first may tend to favor the second, or vice versa. But the linkage cannot be assumed; it needs to be shown. Legal expressivists have failed, as yet, to show a linkage—indeed, I know of no serious attempt to demonstrate that it exists—and for the reasons briefly sketched out here, I am skeptical that any efforts will succeed.¹⁵⁸ Even if Anderson is correct on her own turf (which I do not concede), we have not yet defended a genuine, sentence-

¹⁵⁷ Indeed, Anderson herself acknowledges as much. See ANDERSON, *VALUE IN ETHICS AND ECONOMICS*, *supra* note 143, at 72 ("In private life, one may of course give the interests of the beloved priority over those of strangers. But in public life . . . one may not weight the interests of the beloved more heavily . . .").

¹⁵⁸ Relatedly, legal scholars who show that individual action is expressive, and that this affects what government should do, have not yet given us an expressive theory of governmental action. I have in mind here Gillian K. Hadfield's recent article, *An Expressive Theory of Contract: From Feminist Dilemmas to a Reconceptualization of Rational Choice in Contract Law*, 146 U. PA. L. REV. 1235 (1998). Hadfield claims that certain contractual choices are rationally motivated by expressive considerations and that, where this occurs, the sheer fact of choice does not justify binding the promisor to his or her promise. Rather, government needs some further justification. See *id.* at 1261-82. But this claim does not entail that government itself must respond to meaningful (or meaningless) individual actions by saying something, and indeed Hadfield does not argue that government should do so. See *id.* at 1282 (identifying "reliance interests" and "instrumental justifications . . . flowing from the value of contracting in a given case and the risk that nonenforcement within that category poses to the stability of contracts in other categories" as chief reasons for enforcing expressively-motivated promises). Hadfield is not, by my classification, an expressivist about law or governmental action at all. See also Adler, *Introduction*, *supra* note 17, at 1178-79 (discussing Hadfield's article); Stephen Gardbaum, *Law, Incommensurability, and Expression*, 146 U. PA. L. REV. 1687, 1689-95 (1998) (same).

meaning, expressivist theory of *law*. Let us now consider whether such a theory is, in fact, defensible.

II. EXPRESSIVE THEORIES: THE MAIN EXAMPLES

Part I clarified the concept of an expressive theory. The claim that the linguistic meaning of a legal utterance can have a moral impact is clearly true, but also banal. For example, it is clearly true, but also banal, that the meaning of a legal utterance may cause good or bad outcomes, by virtue of listeners' reactions to what the utterance says. To be sure, it is not uninteresting to specify *which* types of meanings cause *which* outcomes; but the sheer claim that law's meaning can have a causal effect on outcomes hardly produces a genuine expressive theory of law. Various features of legal officials' actions can have a causal effect on outcomes—for example, the time or geographic location at which the actions are performed, or the perceived intelligence of the officials who perform them. Yet we do not think that the correct theory of law is a "temporal" theory, or a "location-based" theory, or an "intelligence-based" theory.

Rather, a genuine expressive theory of law is a theory such that, among the moral factors $F_1 \dots F_n$ constitutive of that theory, at least one refers to the linguistic meaning of legal officials' actions. For at least one factor F_i , it is necessarily the case that meaningless actions fare the same with respect to this factor.¹⁵⁹ Further, I argued in Part I that the only plausible expressive theories of law are sentence-meaning theories, not speaker's-meaning theories. Actions by multimember legal bodies can possess sentence meanings, but typically lack any speaker's meanings beyond whatever sentence meanings they may possess.¹⁶⁰ I will therefore use the term "expressive theory of law" to refer, specifically, to sentence-meaning theories. A genuine sentence-meaning theory of law is a theory such that at least one of its factors refers to the sentence meaning of the actions that legal officials perform. For at least one factor F_i , it is necessarily the case that legal officials' actions lacking a sentence meaning will fare the same with respect to this factor.

Are any such theories correct? Should we adopt one or another expressive (sentence-meaning) theory of law? This Part examines the most prominent expressive theories extant in the scholarly literature and the case law. Section A explores expressive theories of punish-

¹⁵⁹ See *supra* text accompanying notes 130-33.

¹⁶⁰ See *supra* Part I.B.

ment; the problem of punishment has been the locus for the oldest and most sustained debate about legal expressivism. Section B looks at expressive theories of constitutional law, specifically race discrimination and the Establishment Clause. Finally, Section C examines expressive theories of regulation. In all these areas, scholars and jurists have presented and defended moral theories that more or less fit my definition of a genuine expressive (sentence-meaning) theory. My conclusions in this Part are generally skeptical. Neither an expressive theory of punishment, nor of constitutional law, nor of regulation, is ultimately persuasive.

A. *Expressive Theories of Punishment*

Any clearheaded analysis of expressive theories of punishment should distinguish between descriptive and normative claims. I am not the first to make this point,¹⁶¹ but because it is often overlooked, it bears reemphasis.

It may well be true that, as a descriptive matter—as a matter of social facts—there exists a separate institution which we call “punishment” and which is, at least in part, expressive. The central thrust of Joel Feinberg’s oft-cited article, *The Expressive Function of Punishment*, is to argue for this descriptive claim:

Both penalties and punishments are authoritative deprivations for failures; but . . . punishments have an important additional characteristic in common. That characteristic . . . is a certain expressive function: punishment is a conventional device for the expression of attitudes of resentment and indignation, and of judgments of disapproval and reprobation, on the part either of the punishing authority himself or of those “in whose name” the punishment is inflicted. Punishment, in short, has a *symbolic significance* largely missing from other kinds of penalties.¹⁶²

But the fact that the institution we call “punishment” is essentially expressive hardly makes out the normative claim that punishment is justified in virtue of its expressive cast.¹⁶³ Absent some reason to think that social and physical distinctions necessarily mark *moral* distinctions—and surely they do not necessarily do so—it is simply a category

¹⁶¹ See H.L.A. HART, PUNISHMENT AND RESPONSIBILITY 263 (1968); TEN, *supra* note 27, at 42; Primoratz, *supra* note 28, at 187-88. Interestingly, Primoratz subtly overlooks the point in presenting his own, purportedly normative account of punishment. See *infra* text accompanying notes 203-05.

¹⁶² FEINBERG, *supra* note 23, at 97-98.

¹⁶³ Feinberg himself is well aware of this. See *id.* at 98 (distinguishing between the definition of punishment and its justification).

mistake to conclude from Feinberg's descriptive story that a genuine expressive theory of punishment holds true.

Let us define the following practices. *Punishment_H* is defined as the infliction of hard treatment upon an offender.¹⁶⁴ *Punishment_E* is defined as the condemnation of an offender (where condemnation is meant to cover the various kinds of negative, expressive responses that Feinberg describes). *Punishment_{E+H}* is defined as the infliction of hard treatment upon, plus the condemnation of, an offender. Feinberg's descriptive story, if correct, shows that punishment (the institution to which we refer by that name) equals *Punishment_{E+H}*. This hardly shows that, morally, the choice among *Punishment_H*, *Punishment_{E+H}*, and *Punishment_E* is a function or partial function of their expressive properties. For example, "cars" essentially involve internal combustion, "marriage" essentially involves some formal commitment between the parties, and "houses" essentially have walls, but these properties may not have any foundational, moral status—they may not constitute morally basic factors—within the correct normative theories of transportation, friendship, and shelter.¹⁶⁵

What adjudicates the choice between *Punishment_H* on the one hand, and *Punishment_E* or *Punishment_{E+H}* on the other? Why is it morally obligatory, or morally better, or more legitimate, to choose the *E* or *E+H* variants instead of the *H* variant as a response to wrongdoing? *Deterrence* is a standardly invoked criterion by which to evaluate penal responses.¹⁶⁶ One penal response is better than another, qua deterrence, if the first response causes more future would-be wrongdoers to refrain from wrongdoing than the second.¹⁶⁷ Professor Kahan, in his

¹⁶⁴ See *id.* at 95 ("Punishment is defined in effect as the infliction of hard treatment by an authority on a person for his prior failing in some respect (usually an infraction of a rule or command).") (paraphrasing definition of punishment offered by Flew, Benn, and Hart)).

¹⁶⁵ As C.L. Ten puts the point:

It is important, as Hart points out, to distinguish between [a] denunciatory theory which seeks to provide a *justification* of punishment from the different view that it is one of the *defining features* of legal punishment that it expresses the community's condemnation of the offender's act. It is quite consistent to adopt the latter view while believing that the punishment is unjustified and should not be inflicted, or that the justification of punishment is in terms of its deterrent, reformatory, or incapacitative effects.

TEN, *supra* note 27, at 42 (citation omitted).

¹⁶⁶ See, e.g., MICHAEL MOORE, *PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW* 84-88 (1997) (summarizing standardly offered justifications for punishment, including deterrence).

¹⁶⁷ More precisely, one penal response is better than another, qua deterrence, if the first causes more future would-be wrongdoers to refrain, *via the right sort of causal*

recent work on punishment theory, has tried to show that considerations of deterrence justify an expressive response to wrongdoing—either imprisonment (the classic example of *Punishment_{E+H}*) or, in the case of less significant offenses, “shaming” (in effect, *Punishment_E*)—as opposed to nonexpressive alternatives such as fines and community service.¹⁶⁸ As Kahan explains, “[T]he expressive [cast of punishment] might reinforce deterrence . . . through preference formation. The law can discourage criminality not just by ‘raising the cost’ of such behavior through punishments, but also through instilling aversions to the kinds of behavior that the law prohibits.”¹⁶⁹

Kahan goes on to discuss the specific links between public expression of disapproval for some behavior-type and the deterrence of future instances of that behavior: preference-adaptation (“citizens form aversions to the kinds of behavior . . . that the law tells them are unworthy of being valued”¹⁷⁰); belief-dependent preferences (“the willingness of persons to obey various laws is endogenous to their beliefs about whether others view the law as worthy of obedience”¹⁷¹); and goodwill (criminal punishment creates goodwill towards criminal law as a whole, and a concomitant tendency to obey its prohibitions “when criminal punishment confirms, rather than disappoints, shared expectations about what behavior is worthy of moral condemnation”¹⁷²). But none of this, however true and important, advances one bit the expressivist claim that *Punishment_E* or *Punishment_{E+H}* is justified in virtue of what it says. Kahan’s work, insofar as it focuses upon deterrence, is simply an elaborate (and impressive) attempt to show that *Punishment_E* and *Punishment_{E+H}* have a causal impact upon future wrongdoing, rela-

mechanism. Certain causal mechanisms by which penal responses diminish future wrongdoing count as “rehabilitative” or “incapacitative” rather than deterrent. See Garvey, *supra* note 10, at 757 (“Specific deterrence forestalls future offenses by changing the offender’s cost-benefit calculus. Rehabilitation, on the other hand, forestalls future offenses by changing the defendant’s preferences.”). Unless the causal mechanism is rigged so that meaningless actions always fare the same, with respect to that mechanism—and such a mechanism would, then, not count as a plausible specification of what we mean by “deterrence”—nothing I say here about the nonexpressive cast of deterrence depends upon what the specific mechanism is. For example, to use Garvey’s proposal: the meaningless action of executing a criminal versus the meaningless action of confiscating his property can have different impacts on the cost-benefit calculations of future, would-be wrongdoers.

¹⁶⁸ See sources cited *supra* notes 9, 15.

¹⁶⁹ Kahan, *supra* note 9, at 603.

¹⁷⁰ *Id.*

¹⁷¹ *Id.* at 604.

¹⁷² *Id.*; cf. Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453 (1997) (arguing, on purely utilitarian grounds, that the institution of criminal punishment should track collective perceptions of moral desert).

tive to *Punishment_H*.

A deterrent account of expressive penalties, such as Kahan's, is an excellent example of a legal theory that accords moral impact to linguistic meaning but is not a genuine expressive theory. The criterion of deterrence, because it is causal, is necessarily nonexpressive. It asks whether a particular penal response is effective in *causing*¹⁷³ future would-be wrongdoers to refrain from wrongdoing; it does not ask whether a particular penal response *means* something. Two meaningless responses—for example, the linguistically meaningless action of executing a criminal versus the linguistically meaningless action of confiscating his property—can easily have different causal effects on future wrongdoing and thereby fare differently with respect to the factor of deterrence. Deterrence is thus a hopeless place to make out a case for a genuine expressive theory of law.¹⁷⁴

¹⁷³ One subtle objection here is that the right causal theory might itself be expressive. Imagine a theory of human action such that whether *P* performs *A* depends upon a number of factors, including the linguistic meaning of the utterances communicated to *P*. Assume that meaningless actions fare identically with respect to this causal factor so that the factor, and therewith the theory, is genuinely "expressive." Assume further that the causal theory is correct. (The standard theory of human action—the belief/desire theory, namely that *P* tends to perform *A* if *P* desires *O* and *P* believes that *A* leads to *O*—is not "expressive" in this way, but that theory could be wrong.) Imagine, now, that our moral theory evaluates actions in light of their causal upshots. Does not the fact that this moral theory embeds a causal theory that (as it turns out) is expressive make the moral theory itself expressive?

I think not. The causal theory, if it is correct, will not be a one-factor theory. What *P* does surely depends not just on the utterances communicated to him, but also on various other factors, for example, his beliefs or the content of the norms that he accepts. Thus, although meaningless actions will fare identically with respect to the expressive factor incorporated in our causal theory, they need not fare identically, all things considered, in their causal upshots. (One meaningless action by government might lead *P* to do one thing, while another might lead *P* to do something else, depending on the facts *other than* linguistic meaning that bear upon the *nonexpressive* factors incorporated in the causal theory.) What matters, morally, about the meaningless actions are the outcomes that they cause; in that respect, they can differ.

In any event, even if I am incorrect in the claim that a moral theory looking to the causal upshots of action remains nonexpressive notwithstanding the truth of an expressive causal theory, I know of no expressivist who has attempted to defend, let alone succeeded in defending, such a causal theory.

¹⁷⁴ Kahan also claims, in line with a suggestion of Sunstein's, that a deterrent theory is (possibly) "expressive" because it needs a criterion for valuing consequences, which expressivism can provide. See Kahan, *supra* note 9, at 602 ("Without a theory for identifying which outcomes are socially disvalued and how much, it is impossible to know what to deter . . . [O]ne could overtly draw on these [expressive] sensibilities to identify preferred outcomes." (citations omitted)). I respond to this line of argument below, see *infra* note 392.

To be sure, criminal law is not just a matter of deterrence.¹⁷⁵ Retributivism is the standard account as to why a penal response to wrongdoing might be justified, independent of the causal efficacy of that response in preventing future crimes. Retributivism, in turn, comes in different versions, but the simplest and classic version says that moral culpability (wrongdoing plus a culpable mental state on the part of the wrongdoer) is both necessary and sufficient to justify punishment. As Michael Moore explains:

A retributivist punishes because, and only because, the offender deserves it. Retributivism thus stands in stark contrast to utilitarian views that justify punishment of past offenses by the greater good of preventing future offenses. It also contrasts sharply with rehabilitative views, according to which punishment is justified by the reforming good it does the criminal.¹⁷⁶

What, however, does the retributivist mean by "punishment"? Does he or she mean to claim that moral culpability is necessary and sufficient to justify *Punishment_H*? Or, does "punishment" instead refer to some expressive variant, either *Punishment_E* or *Punishment_{E+H}*?

Moore himself, apparently, takes the view that moral culpability justifies *Punishment_H*.¹⁷⁷ Let us call Moore's version of retributivism "Hard Treatment Retributivism." It says that the morally culpable actor deserves *Punishment_H*. Hard Treatment Retributivism is not, of course, a genuine expressive theory, but it is instructive to see why the practice of *Punishment_H* could be mistaken for genuine expression. First, if penal institutions were reliable, the imposition of *Punishment_H* upon an actor would genuinely *signal* that the actor was morally culpable. The actor's hard treatment would, indeed, be "meaningful"—not in the linguistic sense of meaningful, but in the nonlinguistic, or evidentiary, sense that Grice also described.¹⁷⁸ Assuming the reliability

¹⁷⁵ See, e.g., MOORE, *supra* note 166, at 84 ("There is by now a familiar list of prima facie reasons given to justify the institution of punishment. Such a list usually includes: incapacitation, special deterrence, general deterrence, denunciation, rehabilitation, and retribution."). Moore distinguishes between retribution and denunciation because he is not an expressivist about retribution. See *infra* note 177 and accompanying text.

¹⁷⁶ Michael S. Moore, *The Moral Worth of Retribution*, in RESPONSIBILITY, CHARACTER, AND THE EMOTIONS 179, 179 (Ferdinand Schoeman ed., 1987). See MOORE, *supra* note 166, at 104-52, for an updated version of this article. See also *id.* at 83-103, 153-88 (providing a further definition and defense of retributivism).

¹⁷⁷ This seems clear enough from Moore's statement that the "denunciatory," i.e., expressive, account of punishment is a utilitarian rather than retributivist account. See MOORE, *supra* note 166, at 90; Moore, *supra* note 176, at 181.

¹⁷⁸ See *supra* text accompanying notes 70-72.

of the penal institutions, the imposition of *Punishment_H* upon the actor would constitute evidence of her culpability.¹⁷⁹

Second, the actions of the penal officers would possess certain ancillary sentence meanings, extrinsic to the imposition of *Punishment_H*. At a minimum, their actions would possess prescriptive meanings—where, as I explained earlier, “prescriptive meaning” is simply the characteristic meaning of a legal utterance.¹⁸⁰ *Punishment_H* would normally be partly constituted by a package of legal rights, duties, and so forth, and the utterances promulgating these (“The Defendant is hereby required to serve a term of imprisonment, not to exceed . . .”) would be prescriptively meaningful. Note, however, that this communicative feature of state *Punishment_H* does not yet make the case for genuine expressivism—because we have not yet shown that those who impose *Punishment_H* are morally *required* to do so through prescriptively meaningful actions. Indeed, it is hard to see why *Punishment_H* necessitates the creation of fresh legal rights, duties, and the like. Imagine a case in which government punishes a wrongdoer by taking away some of his property—not by depriving him of any rights in the property, but simply by preventing him from gaining physical access to it.¹⁸¹ Why should not this confiscatory but nonprescriptive response, if proportioned to the wrongdoing, constitute a legitimate kind of *Punishment_H* within Hard Treatment Retributivism?

More plausibly, the legitimate imposition of *Punishment_H* by legal officials entails not prescriptive meaning but *justificatory meaning*. The officials will plausibly be required, not merely to have a justification for *Punishment_H* (namely, that the actor was a culpable wrongdoer), but also to *articulate* that justification. So our moral theory, plausibly, says that *Punishment_H* must be accompanied by certain descriptive sentences, explaining the wrong that the actor performed, its proportionality to the hard treatment imposed, and so on. Such a moral theory would, indeed, constitute a genuine expressive theory of law—one which I will discuss at some length later in this Article¹⁸²—but it is

¹⁷⁹ See Michael Davis, *Punishment As Language: Misleading Analogy for Desert Theorists*, 10 LAW & PHIL. 311, 316-18 (1991) (arguing that the justified imposition of hard treatment upon a deserving wrongdoer has nonlinguistic meaning, because it “carries information”).

¹⁸⁰ See *supra* text accompanying notes 104-12.

¹⁸¹ Cf. JOSEPH RAZ, PRACTICAL REASON AND NORMS 157 (Princeton Univ. Press 1990) (1975) (noting that “most sanctions consist in the withdrawal of rights or the imposition of duties” but that “[s]ome sanctions, like capital punishment or whipping, consist in the use of force against a person”).

¹⁸² See *infra* notes 432-39 and accompanying text.

quite different from the expressive theory of *punishment* now extant (in all its variants) in the scholarly literature.¹⁸³ To justify is to utter a sentence conventional for performing a descriptive speech-act, while expressivists about punishment have argued that penal officials should utter sentences conventional for performing nondescriptive speech-acts—sentences of condemnation or denunciation.¹⁸⁴ To quote Feinberg's definition of punishment once more: "[P]unishment is a conventional device for the expression of attitudes of resentment and indignation, and of judgments of disapproval and reprobation."¹⁸⁵ And the requirement that legal officials articulate a justification for *Punishment_H* would, presumably, be but one instance of a generic expressive duty—since justification is plausibly required not only for criminal penalties, but also for a wide array of individual setbacks, including civil penalties and the denial of monetary benefits.¹⁸⁶ Justificatory expressivism would not, *pace* the defenders of an expressive theory of punishment, be specific to, or even exemplified by, the criminal law.

In short, neither the nonlinguistic, signaling function of *Punishment_{HP}* within Hard Treatment Retributivism, nor the generic requirement that legal officials utter sentences apt to communicate a justification for the setbacks they impose, implies that those officials are obliged to condemn or denounce criminal wrongdoing—that wrongdoing should necessarily be followed by *Punishment_E* or *Punishment_{E+H}*. How, then, does one show that? In the remainder of this Section, I will briefly discuss the most recent and prominent attempts within the scholarly literature on punishment to show a necessary connection between wrongdoing and *Punishment_E* or *Punishment_{E+HP}* rather than merely *Punishment_H*. I have in mind the accounts of Rob-

¹⁸³ In particular, it is quite different from the theories summarized and criticized immediately below, namely, the theories of Nozick, Hampton, Primoratz, and Duff. See *infra* text accompanying notes 187-211.

¹⁸⁴ What about Jean Hampton, whose account is discussed below? See *infra* notes 189-202 and accompanying text. Does not she want government to perform some kind of descriptive utterance? Perhaps not; perhaps she wants government to *declare* the wrongdoer and victim to be of equal worth. Cf. *infra* text accompanying note 223 (discussing whether stigmatic meaning is descriptive or declaratory). In any event, if Hampton does want a description, that description is far from being merely justificatory.

¹⁸⁵ FEINBERG, *supra* note 23, at 98.

¹⁸⁶ Cf. Henry J. Friendly, "Some Kind of Hearing," 123 U. PA. L. REV. 1267, 1292 (1975) (suggesting that one element of a full hearing, for procedural due process purposes, is a "statement of reasons" for the decision); Edward L. Rubin, *Due Process and the Administrative State*, 72 CAL. L. REV. 1044 (1984) (describing a wide array of individual setbacks that are accorded procedural due process protection under modern doctrine).

ert Nozick, Jean Hampton, Igor Primoratz, and Anthony Duff.¹⁸⁷ I submit that all of these accounts of punishment expressivism are unpersuasive.

Nozick's account might be called "Expressive Retributivism." While a Hard Treatment Retributivist, such as Michael Moore, claims that the morally culpable actor deserves *Punishment_H*, the Expressive Retributivist argues, instead, that *Punishment_E* or *Punishment_{E+H}* is deserved. The Expressive Retributivist concurs in the use of the concept of *desert* and in that crucial sense is a retributivist, but claims that condemnation, instead of or in addition to hard treatment, is what the wrongdoer deserves. Why might that be the case? Nozick's central idea is that the penal response must *track* the wrongdoing. The simple sequence of immoral action, followed by hard treatment for the actor, is not enough to constitute the genuine punishment that considerations of moral desert require:

If S wrongfully shoots another in a canyon and the sound of the shot causes an avalanche that maims or kills S, then this happens to S because of his wrong act but not because of the wrongness of the act. Since an act's moral qualities, qua moral qualities, seem to lack causal power, *if something is to happen to someone because of the moral quality of his act, this must occur through another's recognition of that moral quality and response to it . . .*

"Poetic justice" involves the wrongdoer's undergoing a consequence that appropriately could be visited upon him in retribution but which was not produced in that way . . .¹⁸⁸

Even if Nozick is correct in the premise that punishment is more than "poetic justice"—that the deserved treatment received by the wrongdoer must be meted out, in virtue of her desert, not just accidentally—it is a large and unwarranted leap from this premise to the conclusion that punishment must be expressive. Imagine a system in which penal officials are trained in Hard Treatment Retributivism and implement that moral view. Does not such a system fulfill the kind of tracking requirement Nozick describes, just as well as a system in which penal officials are trained in Expressive Retributivism and implement *that* moral view? The Nozickean premise is that wrongdoing and penal response must be linked "through another's recognition of that moral quality [wrongdoing]"—thus the emphasized language in the quotation above—but the penal sequence where the official first recognizes

¹⁸⁷ See *supra* note 28 (citing these accounts).

¹⁸⁸ NOZICK, *supra* note 28, at 369-70 (emphasis added).

wrongdoing, then imposes nonexpressive hard treatment for the reason that the treatment is deserved, instantiates quite fully the desired linkage.

A second and quite different attempt to explain why the morally warranted response to wrongdoing is necessarily an expressive response is Jean Hampton's account.¹⁸⁹ Hampton styles herself a retributivist,¹⁹⁰ although this is arguably inaccurate. Retributivism in a broad and fairly unhelpful sense is any theory that takes a penal response to be justified, independent of its effect in reducing future wrongdoing or some other effect valuable within utilitarianism. Retributivism in a more specific sense is exemplified by what I have called Hard Treatment Retributivism and Expressive Retributivism; such views stipulate that *Punishment_H* or *Punishment_E* or *Punishment_{E+H}* is warranted because it is *deserved* by the wrongdoer, independent of any crime-reduction effect or any other positive effect on overall well-being.¹⁹¹ Hampton, unlike Moore and Nozick, is not particularly concerned with desert. Her focus is more on restoring certain components of the *victim's* personhood that are damaged by wrongdoing. Thus, Hampton's theory might plausibly be described as an expressive, compensatory theory of punishment.¹⁹²

What is the link between wrongdoing, compensation, and expression? Hampton argues as follows:

(1) Certain wrongful actions are expressive, insofar as they send a (untrue) message about the victim's lesser moral worth. As Hampton puts it: "[A] person is morally injured when she is the target of behavior whose meaning, appropriately understood by members of the cultural community in which the behavior occurs, represents her value as

¹⁸⁹ See Hampton, *supra* note 28, for the most recent version of Hampton's expressive account of punishment. Earlier versions are presented in Jean Hampton, *An Expressive Theory of Retribution*, in *RETRIBUTIVISM AND ITS CRITICS* 1 (Wesley Cragg ed., 1992); Jean Hampton, *A New Theory of Retribution*, in *LIABILITY AND RESPONSIBILITY* 377 (R.G. Frey & Christopher W. Morris eds., 1991); and JEFFRIE MURPHY & JEAN HAMPTON, *FORGIVENESS AND MERCY* 111-61 (1988).

¹⁹⁰ See Hampton, *supra* note 28, at 1694.

¹⁹¹ See GEORGE SHER, *DESERT* 13 (1987) ("Retributivism without *desert* . . . is like *Hamlet* without the Prince of Denmark.") (quoting Hugo Adam Bedau, *Retribution and the Theory of Punishment*, 75 J. PHIL. 601, 608 (1978)); Moore, *supra* note 176, at 179 ("A retributivist punishes because, and only because, the offender deserves it.").

¹⁹² Indeed, Hampton herself says as much. See Hampton, *supra* note 28, at 1698 ("What these reflections show is that retribution [that is, the expressive response to wrongdoing defended by Hampton] is actually a form of compensation to the victim. . . . [R]etributive justice compensates victims for moral injuries.").

less than the value she should be accorded."¹⁹³ Hampton gives the horrific example of a farmer who reacts to disobedience on the part of his black farmhands by mutilating and burning them.¹⁹⁴

(2) The expressive cast of this kind of wrongdoing is injurious, apart from its simple (nonexpressive) harmfulness, because it "diminishes" the victim. The victim's actual moral worth is not degraded—wrongdoing cannot accomplish that—but an appearance of degradation, what Hampton calls "diminishment,"¹⁹⁵ is indeed the upshot of expressive wrong.

[E]ven when a wrongful action does not inflict a harm, it angers us simply by virtue of what the action says about the person. We care about what people say by their actions because we care about whether our own value, and the value of others, will continue to be respected in our society. The misrepresentation of value implicit in moral injuries . . . threatens to reinforce belief in the wrong theory of value [a theory that takes the victim to be less worthy] by the community.¹⁹⁶

(3) *Punishment_E* or *Punishment_{E+H}* is the appropriate governmental response to the wrongdoing here described, that is, to wrongdoing which means that the victim has lesser moral worth and which in fact diminishes him.¹⁹⁷ Why? Because *Punishment_E* or *Punishment_{E+H}* restores to the victim the appearance of moral worth and thereby reverses the special injury, diminishment, produced by this kind of wrongdoing:

[Punishment] is a response to a wrong that is intended to vindicate the value of the victim denied by the wrongdoer's action through the con-

¹⁹³ *Id.* at 1670.

¹⁹⁴ *See id.* at 1675.

¹⁹⁵ *Id.* at 1673.

¹⁹⁶ *Id.* at 1678; *see id.* at 1671-85 (arguing generally that expressive wrongdoing can diminish a victim and thereby injure him). Hampton actually describes two different ways in which diminishment can injure the victim: by causing "damage to the realization of a victim's value" or "damage to acknowledgment of the victim's value." *Id.* at 1679 (emphasis omitted). My summary of Hampton's view, and the criticism I then offer of that view, focuses on acknowledgment-type damage, but similar criticisms would apply to realization-type damage.

¹⁹⁷ Hampton actually argues for *Punishment_{E+H}* rather than *Punishment_E*, as does Nozick. *See id.* at 1686; NOZICK, *supra* note 28, at 376-77. This feature of their theories is not important for my purposes. If the expressivist shows that morality irreducibly requires *Punishment_E* or *Punishment_{E+H}* then she has indeed made out a genuine expressive theory.

In addition, it bears mention that Hampton does not intend her theory to be limited to, or even exemplified by, governmental responses to wrongdoing, *see* Hampton, *supra* note 28, at 1685, but at a minimum, the theory can be adapted to furnish an expressive theory of (governmental) punishment, as I have done here.

struction of an event that not only repudiates the action's message of superiority over the victim but does so in a way that confirms them as equal by virtue of their humanity.¹⁹⁸

The state is required to respond to expressive wrongdoing by *saying*, "The victim and the wrongdoer are moral equals," and not just by imposing noncommunicative hard treatment upon the wrongdoer.

What is wrong with Hampton's argument? Premises (1) and (2) are problematic—it is problematic that expressive wrongdoing, *Wrongdoing_E*, is really morally distinct from nonexpressive wrongdoing that also harms the victim's status and self-respect—but I will ignore that point for now.¹⁹⁹ Even granting Hampton her first two premises, the inference from those to the conclusion, (3), is unpersuasive. The connection between *Punishment_E* or *Punishment_{E+H}* and the reversal of the victim's diminishment is too weak to make out a noncontingent case for an expressive response to *Wrongdoing_E*. The victim's status within the community and her concomitant self-respect have been lowered or held down; *Wrongdoing_E* has led the community to believe that the victim has lesser moral worth. But it is only contingently true that the best way for government to reverse this status harm is to *communicate* something. For example, government might more effectively achieve equality of status between victim and wrongdoer by coercing the payment of reparations from one to the other, or by bringing it about that the community learns of the victim's virtues and the wrongdoer's flaws, or by conferring upon the victim certain correlates of status, such as wealth and political power. Diminishment, as described by Hampton, is a matter of communal beliefs and norms;²⁰⁰ and what premises (1) and (2) truly imply is not (3), but rather (3'): some governmental action (expressive or not) that best counteracts

¹⁹⁸ Hampton, *supra* note 28, at 1686. See generally *id.* at 1685-98 (arguing for an expressive response to expressive wrongs that diminish victims).

¹⁹⁹ See *infra* text accompanying notes 235-42, 286-90 (distinguishing between the linguistic meaning of an action and its effect upon a person's social status and self-respect).

²⁰⁰ Although Hampton defines diminishment as the appearance of degradation, see Hampton, *supra* note 28, at 1673, such appearance, if injurious to the victim, will surely involve the existence of collective norms according the victim a degraded status or, at a minimum, the existence of shared beliefs that the victim is degraded. This is true, by definition, for what Hampton calls "acknowledgment" damage, but it is also true for what she calls "realization" damage. See *supra* note 196 (articulating these categories). As Hampton notes, "[m]erely publishing a book proclaiming that . . . whites are better than blacks . . . is to deny value and thereby do something morally offensive, but unless other people respond to the book in some way, [the] damage is negligible and society does not bother to respond." Hampton, *supra* note 28, at 1679 (footnotes omitted).

the norms and beliefs constitutive of the victim's diminishment is the appropriate response to *Wrongdoing_E*.²⁰¹

Imagine that by saying, "The victim and the wrongdoer are moral equals," the government in power would actually further degrade the victim's status—for example, because this government is despised within the relevant community. Premises (1) and (2) hardly warrant the conclusion that the despised government is required to utter the harmful sentences, or to say anything at all; yet this is, perversely, what Hampton would oblige the government to do, if she is a genuine expressivist. Alternatively, Hampton could contend for (3') rather than (3), but in that case she has not given us a genuine expressive (sentence-meaning) theory of law. Rather, she has given us a theory that posits the existence of *status-related* goals and constraints covering governmental action. The two kinds of theories are not the same—a point that the case of the despised government is meant to show and to which I will return later, in discussing expressive theories of race discrimination.²⁰²

So much for the expressive theories of Hampton and Nozick. Igor Primoratz has provided a third and quite different argument for the expressive cast of penal responses. Primoratz reasons as follows:

Rules that state standards of behaviour and command categorically imply that actions violating them are wrong, and that such actions are to be condemned, denounced, repudiated. Expressions of this condemnation and repudiation are the index of the validity of the rules and of the acceptance of the conviction that their breaches are wrong in society. If actions of a certain kind can be done without bringing about such a response from society, this indicates that no rule prohibiting such actions is accepted as a valid and binding standard of behaviour.²⁰³

With all due respect to Primoratz—and to Joel Feinberg, who similarly suggests that *Punishment_E* or *Punishment_{E+H}* functions to "vindicate" the legal rules that penalize wrongdoing²⁰⁴—I would suggest that the argument is transparently flawed. Legal rules can surely exist without

²⁰¹ Alan Strudler has pointed out to me that Hampton may want government to respond to *Wrongdoing_E* by correcting the unwarranted reduction of the victim's status and by correcting the unwarranted elevation of the wrongdoer's status, rather than merely doing the first. Construing Hampton's view in this way, however, does not insulate it from the criticism I am raising, namely, that a governmental statement about someone's worth (the victim's, the wrongdoer's, or both) is only contingently linked to her status. I am indebted to Alan for discussions on these issues.

²⁰² See *infra* text accompanying notes 235–42.

²⁰³ Primoratz, *supra* note 28, at 196.

²⁰⁴ See FEINBERG, *supra* note 23, at 104.

the backing of a condemnatory response. Even if conduct-regulating rules must be backed by some sanctions to exist as legal rules,²⁰⁵ it is surely true that these sanctions need not be condemnatory; otherwise, conduct-regulating rules backed by civil sanctions rather than criminal sanctions would not exist as legal rules. If government has a reason to respond to moral wrongdoing by imposing *Punishment_E* or *Punishment_{E,H}* rather than *Punishment_H*, that reason must be the kind of substantive reason elaborated by Nozick and Hampton, and not the false proposition that a rule imposing *Punishment_H* would be legally nonexistent. A second, less obvious flaw in Primoratz's argument is that his point is conceptual or descriptive rather than normative. Assume Primoratzian existence conditions are correct; a deontic statement is not a legal rule unless backed by condemnation. This leaves unanswered the question: Why is government morally required to create legal rules (in the Primoratzian sense)? After all, government might have nonexpressive grounds for creating legal rules (in the Primoratzian sense), e.g., deterrent grounds. The fact that government, for nonexpressive reasons, properly chooses to create such rules does not, in itself, make out an expressive theory. But the simpler objection to Primoratz's argument is that the existence conditions he proposes are just wrong.

Finally, let us turn to the account of punishment proposed by Anthony Duff. Duff's is one of the more recent versions of a moral education theory of punishment,²⁰⁶ and he specifically wants to claim that punishment, insofar as it is educative, is necessarily expressive. Educative theories can focus either on the good of the community—the familiar idea here is that the criminal actor should be rehabilitated so as to prevent future wrongdoing—or on the good of the wrongdoer herself.²⁰⁷ The familiar, rehabilitative theory is clearly not a genuine expressive theory (governmental condemnation of the offender may or may not be the best means to change the preferences and beliefs that led her to commit a crime), and for this and other reasons Duff looks instead to the wrongdoer's own good. "[Punishment] expresses our concern, not merely for some further social benefit or for the restoration of an abstract balance of benefits and burdens, but for the crimi-

²⁰⁵ But see RAZ, *supra* note 181, at 159 (arguing that "resort to sanctions . . . is not a feature which forms part of our concept of law").

²⁰⁶ See Garvey, *supra* note 10, at 763 n.152.

²⁰⁷ See MOORE, *supra* note 166, at 85; see also Garvey, *supra* note 10, at 754-57, 763-65 (distinguishing between the traditional rehabilitative model of punishment and the "educating model").

nal's own good as a rational moral agent and as a member of a moral community.²⁰⁸

This claim is deeply counterintuitive; whatever the purpose of the criminal law, surely it is not (even in part) to confer a benefit upon wrongdoers.²⁰⁹ Even granting Duff his counterintuitive premise, however, we still fail to arrive at expressivism about punishment. Duff argues that *Punishment_{E+H}* educates the wrongdoer and confers a good upon her, because it stimulates her *penitence*:

The purpose of punishment, like that of a penance which is imposed on a recalcitrant sinner, is to bring the criminal to understand the nature and implications of her crime; to repent that crime; and *thus*, by willing her own punishment as a penance which can expiate her crime, to reconcile herself with the Right and with her community.

. . . .

Punishment is, on this conception [*inter alia*] communicative It is communicative in that it seeks to communicate to the criminal a proper understanding of his crime: by imposing on him some material injury which can be seen as injurious even through the eyes of egoistical self-interest, we hope to represent, and to force on his attention, the harm he has done both to others and to himself; by imprisonment . . . we give material and symbolic expression to the spiritual separation created by his crime.²¹⁰

Duff here conflates expressivism about individual action and expressivism about law. It may well be true that penitence itself is essentially expressive. The wrongdoer, *P*, genuinely repents only if she expresses her sorrow, to the victim and the community, as well as experiencing guilt and remorse for her wrong.²¹¹ This hardly shows that the government must bring *P* to repent by condemning her, by expressing sorrow for her, or by uttering sentences additional to the justificatory sentences covered by the generic requirement of public justification. Plausibly, the best way to induce penitence is simply to give the wrongdoer the hard treatment she deserves; to describe that treatment as justified; and thereby to prompt the wrongdoer, herself, first to appreciate her moral culpability and then to act, communicatively, upon that appreciation.

²⁰⁸ DUFF, *supra* note 28, at 234-35.

²⁰⁹ See MOORE, *supra* note 166, at 86-87 ("Such a [theory of punishment] allocates scarce societal resources away from other, more deserving groups that want them (such as retarded and autistic children or the poor) to a group that hardly can be said to deserve such favoured status, and, moreover, does not want such 'benefits.'").

²¹⁰ DUFF, *supra* note 28, at 259-60 (footnotes omitted).

²¹¹ See *id.* at 246-54 (discussing the expressive cast of penitence).

B. *Expressive Theories of Constitutional Law*

1. The Equal Protection Clause

Expressive theorizing about the Equal Protection Clause goes back to its very origins, in *Strauder v. West Virginia*,²¹² the very first race-related case to reach the Court under the Clause. *Strauder* overturned the conviction of a black defendant by a jury from which blacks had been excluded pursuant to an explicitly discriminatory state law. The exclusion was taken to violate equal protection because, in part, of its linguistic content—because it constituted an assertion of black inferiority, or so the Court held. “The very fact that colored people are singled out and expressly denied by a statute all right to participate in the administration of law, as jurors, because of their color . . . is practically a brand upon them, affixed by the law; an assertion of their inferiority”²¹³ The expressivism evident in *Strauder* was carried forward to the modern period by *Brown v. Board of Education*, which described school segregation as causing harm to black schoolchildren in virtue of the fact that “the policy of separating the races is usually interpreted as denoting the inferiority of the negro group,”²¹⁴ and by the constitutional scholars Charles Black and Paul Brest. Professor Black, defending *Brown* against scholarly critics such as Herbert Wechsler, relied substantially on what Black termed the “social meaning” of segregation.²¹⁵ Professor Brest gave that kind of meaning a more specific name, a name now familiar to anyone even passingly acquainted with the literature on equal protection—“stigma”—and developed the seminal scholarly case for an expressivist view of the Equal Protection Clause:

A second and independent rationale for the antidiscrimination principle [embodied by the Equal Protection Clause] is the prevention of the harms which may result from race-dependent decisions. Often, the most obvious harm is the denial of the opportunity to secure a desired benefit—a job, a night’s lodging at a motel, a vote. But this does not completely describe the consequences of race-dependent decision making. Decisions based on assumptions of intrinsic worth and selective indifference inflict psychological injury by stigmatizing their victims as inferior.

²¹² 100 U.S. 303 (1880).

²¹³ *Id.* at 308.

²¹⁴ *Brown v. Board of Educ.*, 347 U.S. 483, 494 (1954) (internal quotation omitted).

²¹⁵ See Black, *supra* note 32, at 426, 427. Wechsler’s famous critique of *Brown* was set forth in Herbert Wechsler, *Toward Neutral Principles of Constitutional Law*, 73 HARV. L. REV. 1, 31-35 (1959).

Moreover, because acts of discrimination tend to occur in pervasive patterns, their victims suffer especially frustrating, cumulative and debilitating injuries.²¹⁶

As this passage suggests, Brest's intent was to map the Equal Protection Clause onto an antidiscrimination principle. That principle, he thought, was supported by certain distinctive moral considerations—in particular, by the fact that state discrimination against blacks had a special linguistic content—that were inapplicable to nondiscriminatory laws which had a disparate impact upon blacks as a group or which otherwise impeded black flourishing.²¹⁷ Such a view, tightly linking stigma with discrimination, is also reflected in the Court's current equal protection jurisprudence, particularly with respect to the legitimacy of affirmative action.²¹⁸ The Court, in defending the doctrinal requirement that all race discrimination (benign or not) be subjected to a "strict scrutiny" test, has relied upon an expressivist account of discrimination: "Classifications based on race carry a danger of stigmatic harm. Unless they are strictly reserved for remedial settings, they may in fact promote notions of racial inferiority and lead to a politics of racial hostility."²¹⁹

Is such a view of the Equal Protection Clause persuasive? First, is the link between discrimination and stigma as tight as Professor Brest and the current Court think? Second, and more deeply, why is it especially wrong for the state to communicate a message of black inferiority? Why are stigmatizing laws worse than meaningless laws that impoverish blacks or that fail to secure them a fair share of resources or welfare?

The first question is connected to a point I made in Part I: that expressivists typically posit a further meaning to legal utterances, beyond the prescriptive meaning those utterances possess just in virtue of defining legal rights, duties, etc.²²⁰ A prescriptive utterance is discriminatory if it includes some reference to the race of the parties

²¹⁶ Brest, *supra* note 32, at 8.

²¹⁷ See *id.* at 6-12 (arguing in favor of the antidiscrimination principle); *id.* at 44-48 (arguing against the view that a law's racially disproportionate impact, as such, is unconstitutional, independent of any link between such impact and past, present, or future discrimination); *id.* at 11 (arguing that the Constitution does not proscribe nondiscriminatory laws that stigmatize blacks or produce "cumulative harms" for individual blacks).

²¹⁸ See *supra* note 37 (citing cases that link stigma with discrimination).

²¹⁹ *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 493 (1989) (citation omitted).

²²⁰ See *supra* Part I.C.

whose rights, duties, etc. are defined.²²¹ For example, the statute at issue in *Strauder* both had prescriptive meaning and was discriminatory. It read as follows: "All white male persons who are twenty-one years of age and who are citizens of this State shall be liable to serve as jurors"²²² But, given such an utterance, it is a further and contingent question whether the utterance is *stigmatic*. This is true however one analyzes the speech-act of "stigmatizing": as a declarative speech-act ("Blacks are hereby declared to be inferior"); as a descriptive speech-act ("Blacks are inferior"); or in some other way.²²³ A discriminatory prescription is stigmatic, in the sentence-meaning sense, only if it triggers the relevant linguistic conventions and thus truly means something like "Blacks are hereby declared to be inferior" or "Blacks are inferior." Did the statute at stake in *Strauder* truly mean that? Was it truly linguistically identical, qua stigma, to a straightforward governmental announcement that blacks "are" or "are hereby declared to be" inferior? Even if this were true of the statutes in *Strauder* and *Brown*, and even if it is typically true, in our racially fractured society, that discriminatory statutes do possess a stigmatic meaning, it is a category mistake to think that an antidiscrimination norm can be reduced to an antistigma norm.

Affirmative action illustrates the point most clearly. As Kenneth Karst explains:

[W]hen government acts to *promote* . . . equal citizenship values . . . , quite a different equal protection issue is presented. Preferential minority admissions to state universities, racial preferences in government hiring, racial preferences aimed at integrating government housing projects—all these differ in the most dramatic way from the purposeful infliction of stigmatic harm.²²⁴

Laws that are drawn along racial lines but provide benefits to blacks need not be stigmatic—and this is true, I suggest, not merely for *justified* affirmative action, but for *unjustified* affirmative action as well. Imagine a local jurisdiction with a majority of black voters that (without good remedial reason) puts in place a system of racial preferences

²²¹ See Adler, *supra* note 47, at 112-21 (discussing concept of discrimination).

²²² 100 U.S. at 305.

²²³ See KOPPELMAN, *supra* note 32, at 67-68 (defining stigma, apparently in declarative terms); *supra* text accompanying notes 99-100 (distinguishing between descriptive and declarative speech-acts). Other stigma theorists, insofar as they construe stigma linguistically, also tend to define it declaratively, as the quotations *supra* at note 75 suggest.

²²⁴ Karst, *Foreword*, *supra* note 32, at 52.

in government employment and contracting.²²⁵ If citizens of the jurisdiction generally understand that the system of preferences is the upshot of ordinary interest-group politics, this time with black employees and contractors gaining at the expense of white employees and contractors, that system (albeit discriminatory and unjustified) presumably lacks the declarative or descriptive sentence meaning constitutive of stigma.²²⁶

This line of argument suggests that expressivism can provide, at most, a partial and incomplete account of the Equal Protection Clause. Intuitively, a racially discriminatory and unjustified law *should* be struck down, whether or not it is stigmatic. At a minimum, placing to one side now the case of affirmative action, it seems intuitively compelling that a racially discriminatory and unjustified law *burdening* blacks should be struck down by the Court, quite independent of whatever (nonprescriptive) linguistic meaning the law possesses. Although it is admittedly difficult to come up with a plausible example of a discriminatory law burdening blacks that is nonstigmatic, it is equally counterintuitive—I suggest—to think that the constitutional validity of a burdensome, discriminatory law should *hinge* upon its (nonprescriptive) linguistic meaning. And there are strong arguments to support this intuition, which I have elsewhere articulated.²²⁷ Race is, in general, morally irrelevant; a law that picks out this morally irrelevant characteristic, stigmatic or not, ought to be invalidated by the Court under the Equal Protection Clause if insufficient grounds exist to justify the discrimination.

The expressivist's best response here is to offer a disjunctive account of the Equal Protection Clause: a law is unconstitutional, under the Clause, if it (1) stigmatizes blacks *or* (2) does not do so, but unjustifiably discriminates on racial lines.²²⁸ On this disjunctive account,

²²⁵ This was how the Supreme Court viewed the affirmative action program struck down in *Croson*. See 488 U.S. at 498-508.

²²⁶ Conversely, a law can be stigmatic without being discriminatory. Charles Lawrence forcefully argues that the nondiscriminatory law upheld by the Court in *Washington v. Davis*, 426 U.S. 229 (1976)—a qualifying test for the D.C. police force that black applicants failed in disproportionate numbers—almost surely sent a message about black inferiority, given the status accorded police officers in our society and the kind of cognitive skills tested. See Lawrence, *supra* note 32, at 369-76.

²²⁷ See Adler, *supra* note 47, at 115-20.

²²⁸ A law's disparate impact upon blacks will not, under current equal protection doctrine, suffice to render the law unconstitutional, see *Washington v. Davis*, 426 U.S. 229 (1976), but it is surely open to argument that current doctrine is, in this respect, misconceived. See, e.g., Owen M. Fiss, *Groups and the Equal Protection Clause*, 5 PHIL. & PUB. AFF. 107, 157 (1976) (arguing that a law or practice that "aggravates . . . the sub-

stigma and discrimination are no longer assumed to track each other perfectly or even particularly well; stigma is seen not as a rationale behind the antidiscrimination rule, but instead as a freestanding criterion for invalidating statutes. As Charles Lawrence has suggested:

[Reviewing courts should] evaluate governmental conduct to see if it conveys a symbolic message to which the culture attaches racial significance. The court would analyze governmental behavior much like a cultural anthropologist might: by considering evidence regarding the historical and social context in which the decision was made and effectuated.²²⁹

But now we reach the second and deeper question I raised above. What is morally distinctive about stigma? Why should this expressive factor be understood as a freestanding component of our best theory of the moral criteria applicable to government action?

Here are what seem to be the standard scholarly moves in answering that question and, in effect, defending an expressivist account of the Equal Protection Clause:

(1) Stigma is tied up with status. Blacks have a second-class status, insofar as the proposition that blacks are inferior is generally believed (or, more stringently, insofar as the treatment of blacks as inferior is supported by a social norm). A stigmatic law maintains or aggravates the second-class status of blacks:

[T]he harms from stigma are not merely psychological. The society also acts toward the stigmatized person on the basis of the stigma. . . . [W]e construct a stigma-theory, an ideology to explain his inferiority and account for the danger he represents. Having done so, we allow ourselves to feel justified in treating the victims of stigma as less than fully human; if they are treated unequally, they are only getting what they "deserve."²³⁰

(2) Status is morally distinct because of its link with self-respect. Self-respect is a component of the good life, a prerequisite for the good life, or both. In any event, self-respect is partly social. A person thought inferior by her fellows cannot possess, or at least predictably will not possess, full self-respect. In short, first-class status is a necessary condition for full self-respect:

ordinate position of a specially disadvantaged group" violates the Equal Protection Clause). The expressivist who thinks that disparate impact (or something like that) constitutes a sufficient ground for unconstitutionality, independent of discrimination or stigma, can simply add a third disjunct to the binary account set forth in the text. The addition of that disjunct, or any additional ones, does not affect the criticisms offered here of the "stigma" disjunct.

²²⁹ Lawrence, *supra* note 32, at 356.

²³⁰ Karst, *Foreword*, *supra* note 32, at 7 (first internal quotations omitted).

When someone is a member of a group that is systematically subordinate to others, and when the group characteristic is highly visible, insults to self-respect are likely to occur nearly every day. An important aspect of a system of caste is that social practices produce a range of obstacles to the development of self-respect, largely because of the presence of the highly visible but morally irrelevant characteristic that gives rise to lower-caste status.²³¹

(3) The moral criterion of equality, constitutionalized in the Fourteenth Amendment, incorporates, as an irreducible component, the principle—call it the anticaste principle, or the antistatutory principle—that all persons shall have first-class status. Stigma is therefore morally and constitutionally distinct:

Stigmatization is the process by which the dominant group in society differentiates itself from others by setting them apart, treating them as less than fully human, denying them acceptance by the organized community, and excluding them from participating in that community as equals. If the equal protection clause guarantees the right to be treated as an equal, the constitutional claim in question can be reduced to [or at least contain as an irreducible component] a claim to be free from stigma.²³²

This general line of argument has been developed by Brest himself and, subsequently, by a number of the leading scholars on the constitutional status of race, particularly Kenneth Karst, Charles Lawrence, Cass Sunstein, and, most recently and extensively, Andrew Koppelman.²³³ These scholars do not necessarily share Brest's view that stigma and discrimination are tightly linked, but they do all seem

²³¹ Sunstein, *supra* note 32, at 2430. Professor Koppelman articulates the claim just as succinctly:

The most obvious harm caused by stigma is that of internalized self-hatred, [a] degenerating sense of nobodiness . . . [H]uman beings . . . whose daily experience tells them that almost nowhere in society are they respected and granted the ordinary dignity and courtesy accorded to others will, as a matter of course, begin to doubt their own worth.

KOPPELMAN, *supra* note 32, at 61 (internal quotations omitted).

²³² Lawrence, *supra* note 32, at 350 (footnotes and internal quotations omitted). To quote Professor Sunstein, whose article on the subject is entitled *The Anticaste Principle*:

[I]n American constitutional law, an important equality principle stems from opposition to caste. . . . The controlling principle is that no group may be made into second-class citizens. Instead of asking "Are blacks or women similarly situated to whites or men, and if so have they been treated differently?" we should ask "Does the law or practice in question contribute to the maintenance of second-class citizenship, or lower-caste status, for blacks or women?"

Sunstein, *supra* note 32, at 2428-29.

²³³ See *supra* note 32.

to agree with him that stigma is distinctively wrong as a moral and constitutional matter. Indeed, that proposition has become a shibboleth in the academic literature.²³⁴

So where does the argument just sketched go awry? It goes awry, I suggest, at the same juncture that Jean Hampton's expressive theory of punishment does:²³⁵ in equating stigma with status harm or, more generally, in equating the meaning of governmental decisions with their cultural impact. Although Hampton uses esoteric terms like "moral injury" and "diminishment" instead of "status" and "second-class citizen," her focus is really the same as that of the stigma theorists: on the harm to status and, concomitantly, self-respect caused by certain acts of wrongdoing.²³⁶ But it is only contingently true that a statement by the government ("The victim and wrongdoer are moral equals") will correct that type of harm; similarly, it is only contingently true that a statement by the government ("Blacks are inferior to whites" or "Blacks are hereby declared to be inferior to whites") will cause it. A governmental decision is stigmatic if it invokes the right kind of linguistic convention, specifically a descriptive or declarative one. But the utterance of the appropriate description or declaration, even by government, is neither necessary nor sufficient for status harm. In short, constitutional scholars such as Brest, Lawrence, Karst, Koppelman, and Sunstein have given us a powerful, *cultural* theory of the Equal Protection Clause,²³⁷ but the cultural and moral factor that these scholars make salient—status—is not an expressive factor in the sense delineated in Part I.

The proposition that a governmental action can cause status harm to blacks without stigmatizing them is, I think, fairly hard to deny. Richard McAdams, one of the leading norms scholars, has argued at length and with rigor that private race discrimination is a mechanism by which whites enhance their status at the expense of blacks.²³⁸ McAdams's work is relevant here because it implies that a legal prohibition on private discrimination will elevate black status quite independent of what that prohibition itself means. For example, a prohi-

²³⁴ See, e.g., LAURENCE H. TRIBE, *AMERICAN CONSTITUTIONAL LAW* 1477 (2d ed. 1988) (arguing that the stigma rationale for the decision in *Brown v. Board of Education* is the "most obvious" and "most persuasive").

²³⁵ See *supra* text accompanying notes 189-202 (criticizing Hampton).

²³⁶ See Hampton, *supra* note 28, at 1671-85.

²³⁷ See KOPPELMAN, *supra* note 32, at 136 ("[T]he purpose of antidiscrimination law is to help destroy the cultural processes by which certain groups are constituted as having an inferior status.").

²³⁸ See Richard H. McAdams, *Cooperation and Conflict*, *supra* note 44.

bition can elevate black status by improving the economic position of blacks or by forcing whites to come into contact with blacks, thereby disabusing whites of certain false beliefs.²³⁹ Conversely, a state's repeal of antidiscrimination laws (or its failure to enact such laws in the first place) would cause status harm to blacks, even if the repeal or nonenactment symbolized nothing. Another plausible example of status impact without linguistic content is affirmative action. A standard defense of affirmative action programs is that they increase the wealth, power, and professional standing of individual black beneficiaries, thereby improving the status of blacks as a group.²⁴⁰ The idea here is that dampening racial skewness in certain indicia or components of status, namely wealth, power, and professional standing, will ameliorate status differences between racial groups; there need not be a further claim that governmental decisions changing the skewness are *themselves* meaningful.

The converse proposition—that governmental action can stigmatize without causing status harm—is, at first, much harder to accept. I should note that the truth of this converse proposition is not crucial for my case against expressivism. As long as stigma, albeit a sufficient condition for status harm, is not a necessary condition for status harm—as long as the set of status-harmful statutes includes some *meaningless* statutes, such as the meaningless statutes mentioned in the preceding paragraph—a moral theory of the Equal Protection Clause grounded ultimately on status and self-respect will not be a genuine expressive theory. What matters, within such a theory, is the effect of governmental action on certain collective beliefs and norms. Meaningless actions can fare differently with respect to the specified beliefs and norms, and the most perspicuous description of a governmental action, within this theory, will refer just to its cultural effect, not to its linguistic meaning—for a linguistic description would miss the case of an action that is linguistically meaningless but has a negative cultural impact, namely, lowering the status of blacks, and thereby causes

²³⁹ See *id.* at 1081.

In sum, law may correct the market failure of discrimination in three ways: by raising the costs and lowering the productive returns of certain forms of subordination; by increasing the racial diversity of socially connected groups; . . . and by symbolizing a consensus that the rationalizations for the subordination strategy are, in fact, mere rationalizations.

Id. Notably, only the last of McAdams's three mechanisms attaches a symbolic content to antidiscrimination law itself.

²⁴⁰ For a recent statement of this view, see Paul Brest & Miranda Oshige, *Affirmative Action for Whom?*, 47 STAN. L. REV. 855, 867-72 (1995).

status harm.

In any event, I suggest that the converse proposition *is* correct. Stigma is not even sufficient, let alone necessary, for status harm. Imagine a weakened government that, in the vain hope of appealing to white voters, begins to play the "race card" by sending subtle and not-so-subtle messages of black inferiority. Do we want to say that such a government inevitably exacerbates or reinforces a system of racial castes? Presumably, the cultural impact of those messages is contingent on additional factors, such as whether white voters are even listening to what the weakened officials are trying to tell them. The voters may be too alienated from politics to listen or to care; they may discount either all governmental messages or all messages from this particular government.

The distinction I am emphasizing here, between cultural impact and linguistic meaning, is obscured in the scholarly literature on equal protection given a persistent equivocation about the definition of "stigma." Sometimes, "stigma" is *defined* in terms of cultural impact: "Systematic differences [between whites and blacks] help produce frequent injuries to self respect—the time-honored constitutional notion of 'stigma.' . . . [S]tigma is part of what it means to be a member of a lower caste."²⁴¹ On this definition, a stigmatic law is sufficient as well as necessary for status harm, but the conceptual link between stigma and language is severed. Conversely, if stigma is defined in linguistic terms—if a governmental action is "stigmatic" just insofar as it possesses a particular descriptive or declarative meaning—then "stigma" is neither necessary nor sufficient for status harm, and a stigma-based account of the Equal Protection Clause, albeit a genuine expressive account, is unpersuasive.²⁴²

²⁴¹ Sunstein, *supra* note 32, at 2430, 2432; see also Karst, *Foreword*, *supra* note 32, at 49-53 (suggesting that stigma is better conceptualized in terms of cultural impact rather than in terms of "purposeful stigmatizing action" by government).

²⁴² Might the theory presented and criticized in this subsection, as well as the parallel theory developed in the Establishment Clause context, see *infra* text accompanying notes 281-285, be modified such that stigma in the linguistic sense is linked directly with self-respect—rather than indirectly, via status—and thereby be rendered immune to the objections I have advanced? Would that not be a viable and genuinely expressive (linguistic) theory? It would be genuinely (linguistically) expressive, but it would not, I think, be viable. The described modification would, indeed, make irrelevant my point that governmental stigma and social status are only contingently linked; but it would involve the implausible claim that a person's self-respect is necessarily degraded by a linguistically stigmatic governmental utterance even if his status remains first-class (say, because the uttering officials are despised or ignored). It is the implausibility of this claim that, I think, has motivated expressivists about the Equal Protection Clause to build status into their accounts, as an intermediate link. I am indebted to Seth Kre-

* *

The arguments I have developed in this subsection leave open the possibility that stigma (in the linguistic sense) might play a *doctrinal* role within the doctrines implemented by constitutional reviewing courts pursuant to the Equal Protection Clause. Perhaps all of the following hold true: (1) the linguistic meaning of a governmental action is not morally relevant, *per se*, within our best moral theory of the Equal Protection Clause; (2) the effect of a governmental action upon black status and self-respect is morally relevant, *per se*, within our best moral theory of the Equal Protection Clause; (3) a cultural impact test would be very difficult for reviewing courts to administer; (4) stigma tracks cultural impact reasonably well, and a stigma test would be easier to administer than a cultural impact test; therefore, (5) a stigma test should be part of Equal Protection Clause doctrine.²⁴³ In general, legal doctrines can refer to the linguistic content of the activities that they cover, whether or not the moral factors governing and animating those doctrines refer to linguistic content. Law need not directly incorporate morality.²⁴⁴

But I would not count the case where nonexpressive morality leads to expressive doctrine (constitutional or otherwise) as a case of genuine expressivism. For example, a preference-utilitarian might support laws prohibiting "libel" or "sedition" or "obscenity," but that hardly makes her an expressivist. "Expressivism," as I use the term, marks out something special about morality itself, not about the host

imer for clarifying discussions on these issues.

Does my view of the Equal Protection Clause mean that government has no reason not to stigmatize when the stigma does not cause status harm? No. Sometimes, stigmatic laws will be discriminatory. And when they are not, government will still have a reason (although not a constitutional reason) to refrain from stigma: either that the stigmatic utterance is simply unjustified (for presumably such an utterance, status-harmful or not, is morally suboptimal in one way or another) or that the stigmatic utterance constitutes a lie and therefore implicates the no-lying constraint. *See infra* text accompanying notes 404-15 (discussing this constraint).

²⁴³ Among contemporary legal scholars, Frederick Schauer is well known for his forceful and prolific analysis of the way in which moral reason can and does obtain for legal doctrines *not* to directly incorporate moral criteria—for example, because actors have limited epistemic abilities, because they cannot be fully trusted, because of the exigencies of coordination, and so on. *See, e.g.*, FREDERICK SCHAUER, *PLAYING BY THE RULES: A PHILOSOPHICAL EXAMINATION OF RULE-BASED DECISION-MAKING IN LAW AND IN LIFE* (1991). Relatedly, there is now a well-accepted distinction in the philosophical literature between moral criteria and morally justified decision procedures. *See, e.g.*, DAVID O. BRINK, *MORAL REALISM AND THE FOUNDATIONS OF ETHICS* 216-17 (1989).

²⁴⁴ At least on a positivist account of the nature of law. *See supra* note 106 (distinguishing between positivist and natural law accounts, and citing sources).

of doctrines, practices, and institutions that our moral view will then inform. It is fair to use the term in this special, restricted sense—not only because many expressivists have used it thus, but also, and more importantly, because the task of developing a moral view is logically and practically *prior* to the task of evaluating doctrines, institutions, and practices. We need, first, to know what the criteria for evaluation are—specifically, whether those criteria, or at least some of them, make essential reference to linguistic meaning. I have argued that the moral criteria or factors animating the Equal Protection Clause do not thus refer. At bottom, a law violates constitutional equality by virtue of the fact that it unjustifiably discriminates on racial lines, by virtue of the fact that it causes status harm to blacks, or perhaps by virtue of other nonexpressive properties (for example, its disparate impact on blacks),²⁴⁵ but not by virtue of the fact that it possess a particular descriptive or declarative sentence meaning.

2. The Establishment Clause

A second area of constitutional law and scholarship where expressivism has been particularly important is the Establishment Clause. In recent years, the symbolic content of governmental action challenged under this clause—whether the action constitutes an “endorsement” of a particular religion or of religion in general—has come to play a significant role in Supreme Court doctrine. This focus on symbolism stems from Justice O’Connor’s concurring opinion in *Lynch v. Donnelly*, where she proposed a modification to the then-standard test for Establishment Clause cases, the so-called “*Lemon* test.”²⁴⁶ In *Lemon v. Kurtzman*, the Court held that the proper standard was a three-pronged test, requiring that “[f]irst, the statute must have a secular legislative purpose; second, its principal or primary effect must be one that neither advances nor inhibits religion; [and] finally, the statute must not foster ‘an excessive government entanglement with religion.’”²⁴⁷ Justice O’Connor suggested that the “purpose” prong of *Lemon* should be modified to ask whether government intended to communicate a message of endorsement or disapproval and that the “effect” prong should be modified to ask whether such a message was, in fact, sent:

The meaning of a statement to its audience depends both on the in-

²⁴⁵ See *supra* note 228.

²⁴⁶ 465 U.S. 668, 687-94 (1984) (O’Connor, J., concurring).

²⁴⁷ 403 U.S. 602, 612-13 (1971) (citation omitted).

tention of the speaker and on the "objective" meaning of the statement in the community. Some listeners need not rely solely on the words themselves in discerning the speaker's intent: they can judge the intent by, for example, examining the context of the statement or asking questions of the speaker. Other listeners do not have or will not seek access to such evidence of intent. They will rely instead on the words themselves; for them the message actually conveyed may be something not actually intended. . . .

The purpose prong of the *Lemon* test asks whether government's actual purpose is to endorse or disapprove of religion. The effect prong asks whether, irrespective of government's actual purpose, the practice under review in fact conveys a message of endorsement or disapproval. An affirmative answer to either question should render the challenged practice invalid.²⁴⁸

O'Connor's proposal here maps nicely onto the two Gricean categories of linguistic meaning: speaker's meaning and sentence meaning.²⁴⁹ What she, in effect, proposed in *Lynch* was that governmental action endorsing religion, in either the speaker's-meaning or the sentence-meaning sense of "endorsement," should be held unconstitutional under the Establishment Clause. Soon thereafter, in *School District v. Ball*, the Court itself incorporated "endorsement" into its Establishment Clause analysis—stating that "an important concern of the [*Lemon*] effect test is whether the symbolic union of church and state effected by the challenged governmental action is sufficiently likely to be perceived by adherents of the controlling denominations as an endorsement, and by the nonadherents as a disapproval, of their individual religious choices"²⁵⁰—and has done so numerous times since.²⁵¹ Most recently, in *Capitol Square Review and Advisory Board v. Pinette*,²⁵² both the plurality opinion²⁵³ and all of the dissenting and concurring justices²⁵⁴ argued for some kind of "endorsement" analysis

²⁴⁸ *Lynch*, 465 U.S. at 690 (O'Connor, J., concurring).

²⁴⁹ See *supra* Part I.B (explicating these two kinds of linguistic meaning).

²⁵⁰ 473 U.S. 373, 390 (1985); see also *id.* at 389 (stating that if government "identification [with religion] conveys a message of government endorsement or disapproval of religion, a core purpose of the Establishment Clause is violated" (citing *Lynch*, 465 U.S. at 688 (O'Connor, J., concurring))).

²⁵¹ For summaries of the doctrinal evolution of the endorsement analysis, with citations to the Supreme Court case law, see Epstein, *supra* note 117, at 2124-37; and Kent Greenawalt, *Quo Vadis: The Status and Prospects of "Tests" Under the Religion Clauses*, 1995 SUP. CT. REV. 323, 359-80.

²⁵² 515 U.S. 753 (1995).

²⁵³ The portion of Justice Scalia's opinion that discussed endorsement was for a plurality of the Court; it was joined by Chief Justice Rehnquist and Justices Kennedy and Thomas. See *id.* at 763-70.

²⁵⁴ See *id.* at 772-83 (O'Connor, J., concurring in part and concurring in the judg-

within Establishment Clause doctrine.²⁵⁵

As *Pinette* reveals, the precise doctrinal function of "endorsement" remains uncertain. This uncertainty is connected to the larger uncertainty about the fate of *Lemon*. A majority of justices currently sitting have criticized the *Lemon* test—the Court in recent years has regularly failed to invoke it—but the test has not been officially overruled or replaced, and no majority of the Court has yet coalesced on a replacement test.²⁵⁶ Nonetheless, if *Pinette* is any indication, "endorsement" will continue to figure in the Court's current, unsettled jurisprudence and will likely be part of whatever official replacement for *Lemon* eventually emerges.

Does this make sense? Why should the constitutionality of legal officials' actions under the Establishment Clause be a function of what the actions mean? My focus here will be on the sentence-meaning, not the speaker's-meaning construal of "endorsement." In Part I of this Article, I adduced some general considerations against speaker's-meaning theories of law,²⁵⁷ and I suggest that those considerations hold good here. *Pace* Justice O'Connor, official actions (at least the actions usually at stake in Establishment Clause cases, those of legislatures, city councils, or other multimember bodies) have no speaker's meaning beyond whatever sentence meaning they might have. Professor Steven Smith has elaborated this and other criticisms of the speaker's-meaning component of Justice O'Connor's proposal in *Lynch*.²⁵⁸ I will therefore train my attention on the sentence-meaning component.

Note, to begin, that an expressive or linguistic test can function, at best, as only one portion of Establishment Clause analysis, not as the whole of it. *Lynch* suggests to the contrary,²⁵⁹ as have some constitu-

ment); *id.* at 797-815 (Stevens, J., dissenting); *id.* at 817-18 (Ginsburg, J., dissenting). See generally Greenawalt, *supra* note 251, at 370-75.

²⁵⁵ See also *Agostini v. Felton*, 521 U.S. 203, 227-28, 235 (1997) (considering whether an educational program, challenged under the Establishment Clause, constitutes a "symbolic union" between government and religion, and concluding that it cannot be viewed as an "endorsement of religion").

²⁵⁶ See Greenawalt, *supra* note 251, at 323-29, 359-61 (summarizing the demise of *Lemon*).

²⁵⁷ See *supra* text accompanying notes 83-89.

²⁵⁸ See Steven D. Smith, *Symbols, Perceptions, and Doctrinal Illusions: Establishment Neutrality and the "No Endorsement" Test*, 86 MICH. L. REV. 266, 283-91 (1987).

²⁵⁹ Specifically, Justice O'Connor's opinion in *Lynch* suggests that all but the no-"entanglement" aspect of Establishment Clause doctrine can be reduced to an endorsement analysis. See *Lynch*, 465 U.S. at 687-89.

tional scholars,²⁶⁰ but the suggestion seems quite implausible. A parallel problem was noted above in our discussion of the "stigma" interpretation of the Equal Protection Clause—some discriminatory and thereby unconstitutional laws might be nonstigmatic²⁶¹—and, in the Establishment Clause context, the problem is considerably more serious and obvious. A wide variety of laws can surely constitute "the establishment of religion" without being linguistically equivalent (in the sentence-meaning sense) to the utterances "Government endorses Religion *R*" or "Government endorses religion in general."

Lynch itself involved the display of an undeniably symbolic object, a crèche, by the government.²⁶² Assume that, in this kind of case, the constitutionality of the governmental action is indeed a function of its semantics.²⁶³ Even so, such cases are only a fairly small portion of the Establishment Clause case law.²⁶⁴ A much larger portion consists of cases involving aid to religious institutions.²⁶⁵ If cases about accom-

²⁶⁰ See Donald L. Beschle, *The Conservative as Liberal: The Religion Clauses, Liberal Neutrality, and the Approach of Justice O'Connor*, 62 NOTRE DAME L. REV. 151, 171-82 (1987) (arguing that "neutrality" should be the basis for Establishment Clause analysis and construing neutrality as nonendorsement); Loewy, *supra* note 76, at 1049-52 (supporting Justice O'Connor's proposed reformulation of the *Lemon* test in terms of endorsement); William P. Marshall, "We Know It When We See It": *The Supreme Court and Establishment*, 59 S. CAL. L. REV. 495 (1986) (arguing that Establishment Clause cases cannot be decided in light of objective principles, given inherent tensions in the area, and that establishment should therefore be generally understood in symbolic terms).

²⁶¹ See *supra* text accompanying notes 220-27.

²⁶² See *Lynch*, 465 U.S. at 671-72.

²⁶³ See Greenawalt, *supra* note 251, at 370 (stating that "a reading of [*Pinette*] suggests that all Justices agree that endorsement is the proper inquiry if the government itself erects a crèche, cross, or menorah on its property").

²⁶⁴ Besides *Lynch* and *Pinette*, the only other Establishment Clause case involving the display of a symbolic object by the government or on government property (as in *Pinette*) is *County of Allegheny v. ACLU*, 492 U.S. 573 (1989). Even if one (plausibly) includes the public prayer cases in the "symbolic" category along with cases like *Lynch*, *Pinette*, and *Allegheny*, such that endorsement is both necessary and sufficient for an Establishment Clause violation within this category, the category is still much smaller (in sheer number of cases at the Supreme Court level) than the school aid cases, let alone the totality of the remaining Establishment Clause cases. Obviously, the proportion of Supreme Court cases in a particular category is not conclusive evidence of the proportion of cases in that category in the lower courts or the proportion of Establishment Clause controversies overall (adjudicated or not) within the category; but it is good evidence, absent a reason to think that the Supreme Court sample is skewed. In this instance, I am aware of no reason to think that "symbolic" cases should constitute a significantly larger portion of the lower-court cases, or of Establishment Clause controversies overall, than of the Court's docket.

²⁶⁵ "Aid" here means aid, financial or other, that is not purely symbolic, such as a grant of textbooks or of government-employed teachers' services. See, e.g., *Agostini v. Felton*, 521 U.S. 203 (1997); *Zobrest v. Catalina Foothills Sch. Dist.*, 509 U.S. 1 (1993);

modation for religious actors,²⁶⁶ cases about public prayer,²⁶⁷ cases about the delegation of power to religious bodies,²⁶⁸ and cases about the need for government officials to make decisions concerning religious doctrine²⁶⁹ are also included, the power of the "endorsement" test seems yet smaller. I suggest that, in all of these areas, on *any* plausible view of the Establishment Clause—whether a separationist view, a neutralist view, an accommodationist view, or a view focused on governmental coercion²⁷⁰—government actions that lack the sentence meaning of "endorsement" might nonetheless be unconstitutional. As Michael McConnell puts it:

Strict separationists will take the position that any provision of financial or other assistance to religion is an endorsement. Advocates of "facial neutrality" will take the position that any action that "singles out" religion for special treatment . . . is an endorsement. Accommodationists will say that benefits to religion that are either facially neutral or that

Bowen v. Kendrick, 487 U.S. 589 (1988); Witters v. Washington Dep't of Servs. for the Blind, 474 U.S. 481 (1986); Grand Rapids Sch. Dist. v. Ball, 473 U.S. 373 (1985); Mueller v. Allen, 463 U.S. 388 (1983); Committee for Pub. Educ. and Religious Liberty v. Regan, 444 U.S. 646 (1980); Wolman v. Walter, 433 U.S. 229 (1977); Meek v. Pittenger, 421 U.S. 349 (1975); Committee for Pub. Educ. and Religious Liberty v. Nyquist, 413 U.S. 756 (1973); Hunt v. McNair, 413 U.S. 734 (1973); Levitt v. Committee for Pub. Educ. and Religious Liberty, 413 U.S. 472 (1973); Tilton v. Richardson, 403 U.S. 672 (1971); Lemon v. Kurtzman, 403 U.S. 602 (1971); Board of Educ. v. Allen, 392 U.S. 236 (1968).

²⁶⁶ See, e.g., Board of Educ. v. Mergens, 496 U.S. 226 (1990); Texas Monthly, Inc. v. Bullock, 489 U.S. 1 (1989); Corporation of Presiding Bishop v. Amos, 483 U.S. 327 (1987); Estate of Thornton v. Caldor, Inc., 472 U.S. 703 (1985); Walz v. Tax Comm'n of New York, 397 U.S. 664 (1970); Zorach v. Clauson, 343 U.S. 306 (1952).

²⁶⁷ By "public prayer" I mean the use of prayer or other religious texts in schools, legislatures, or other public fora. See, e.g., Lee v. Weisman, 505 U.S. 577 (1992); Wallace v. Jaffree, 472 U.S. 38 (1985); Marsh v. Chambers, 463 U.S. 783 (1983); Stone v. Graham, 449 U.S. 39 (1980); Abington Sch. Dist. v. Schempp, 374 U.S. 203 (1963); Engel v. Vitale, 370 U.S. 421 (1962). The cases involving religion in the school curriculum might be put in or near this category. See, e.g., Edwards v. Aguillard, 482 U.S. 578 (1987); Epperson v. Arkansas, 393 U.S. 97 (1968).

²⁶⁸ See, e.g., Board of Educ. v. Grumet, 512 U.S. 687 (1994); Larkin v. Grendel's Den, Inc., 459 U.S. 116 (1982).

²⁶⁹ See, e.g., Serbian E. Orthodox Diocese v. Milivojevic, 426 U.S. 696 (1976).

²⁷⁰ Good exemplars of these views are Philip B. Kurland, *Of Church and State and the Supreme Court*, 29 U. CHI. L. REV. 1 (1961) (neutralism—specifically, a defense of facial neutrality); Douglas Laycock, *Formal, Substantive, and Disaggregated Neutrality Toward Religion*, 39 DEPAUL L. REV. 993 (1990) (neutralism); Michael W. McConnell, *Accommodation of Religion*, 1985 SUP. CT. REV. 1 (accommodationism); Michael Stokes Paulsen, *Lemon is Dead*, 43 CASE W. RES. L. REV. 795 (1993) (coercion); and Kathleen M. Sullivan, *Religion and Liberal Democracy*, 59 U. CHI. L. REV. 195 (1992) (separationism). The line between certain variants of neutralism, such as Laycock's, and accommodationism is a thin one.

accommodate the free exercise of religion are neutral in their symbolic effect, and that anything less would be an expression of disapproval.²⁷¹

I do not think McConnell is right that the concept of endorsement is *necessarily* parasitic on nonexpressive accounts of the Establishment Clause. Government “endorses” religion in the linguistic or expressive sense when it invokes the relevant linguistic convention for performing the speech-act of endorsement: when it utters a linguistic statement the content of which is “we endorse religion” or “we support religion” or something like that. But I do agree with McConnell that any plausible view of the Establishment Clause will include factors such as separation, coercion, neutrality, or accommodation that cannot be reduced to linguistic meaning.²⁷²

To see the point, consider the neutralist defense of the endorsement test offered by Professor Donald Beschle. Beschle argues as follows: “The essential core of liberal neutrality as applied to the religion clauses . . . should be the principle that government may not endorse one set of religious beliefs over another, endorse religion over irreligion, or irreligion over religion.”²⁷³ Does Beschle really mean this? Imagine a spending measure that grants funds directly to certain churches, identified by name, and only to those churches. Surely this law is nonneutral,²⁷⁴ and it seems that Beschle himself would concede

²⁷¹ McConnell, *supra* note 104, at 150-51.

²⁷² The endorsement concept is not necessarily parasitic on nonexpressive accounts of the Establishment Clause, but it is *possibly* parasitic—it can be formulated so that state action endorses religion if and only if that action is nonexpressively unjustified—and perhaps McConnell is correct that the formulations influential within the Supreme Court case law are indeed parasitic in this way. Cf. *Wallace v. Jaffree*, 472 U.S. 38, 76, 83 (1985) (O'Connor, J., concurring in the judgment) (asserting that “[t]he relevant issue is whether an objective observer, acquainted with the text, legislative history, and implementation of the statute, would perceive it as a state endorsement of [religion]” where such an observer is, *inter alia*, “assume[d] . . . [to be] acquainted with the Free Exercise Clause and the values it promotes”).

²⁷³ Beschle, *supra* note 260, at 182.

²⁷⁴ More precisely, the law is nonneutral apart from the special case where the church names are good proxies for nonreligious characteristics, such as the characteristic of being particularly well-suited to provide certain services. See McConnell, *supra* note 104, at 144-45 (noting that “Congress made a large grant to the Roman Catholic Church a few years ago for the purpose of assisting illegal aliens in applying for amnesty . . . [because] the Catholic Church is uniquely positioned to reach them”). The accommodationist might object that *religious* characteristics can themselves be foundationally relevant, rather than mere proxies—a law that singles out and thereby accommodates religious practices can be justifiably targeted at religious liberty and nothing else, or so the accommodationist will argue—but it is hard to see how a law granting funds only to certain churches could be a justifiable accommodation and in that sense “neutral.” See *id.* at 185 (“[T]he taxpayer has a right to insist that the government not

as much, for he says elsewhere that "aid to only one or any number of chosen religions to the exclusion of others would be impermissible."²⁷⁵ But it is an open question whether the law is linguistically equivalent to a statement by government, "We endorse religion *R*," or to any other (nonprescriptive) statement at all. The churches might want material aid rather than endorsement,²⁷⁶ and the legislature might either frankly admit as much or make it clear from context that this was simply pork-barrel politics.²⁷⁷

The better account of neutrality—the one that explains why the spending measure, despite its meaninglessness, violates the Establishment Clause—is not an expressive account, but a *justificatory* account. Liberal neutrality in general constrains the types of moral or evaluative propositions with reference to which government decisions can be defended, classically propositions about the relative goodness of ways of life.²⁷⁸ Liberal neutrality in matters of religion means, specifically, that propositions such as "Religion *R* is the one true religion" or "Religion *R* constitutes a better way of life than Religion *S*" are unavailable to justify government decisions. The liberal neutralist in the justificatory sense will, to be sure, count a government endorsement of religion as nonneutral and unconstitutional; but she will also count meaningless action as unconstitutional just insofar as that action is indefensible relative to a suitably constrained set of moral propositions. Liberal neutralism in the justificatory sense is, therefore, not a genuine expressive theory of law. Beschle's neutralism is, but it is an implausible variant of neutralism (as the spending example is meant to suggest).

give tax dollars to religion *qua* religion, or in a way that favors . . . one religion over another.").

²⁷⁵ Beschle, *supra* note 260, at 184.

²⁷⁶ Thus Smith comments:

To be sure, some politicians *do* seek to attract the votes of religious persons by sending messages endorsing religion. But other supporters of governmental assistance to religious interests or institutions may often attempt, if only for tactical reasons, to provide such assistance in disguised forms carefully crafted to *avoid* endorsing religion . . . [I]t would be disingenuous for the supporters of such measures to deny that they hope to *help*, or to *advance*, religion. But their concern—and their intent in backing such measures—is apparently to *extend material assistance* to religion, not to send messages *endorsing* religion.

Smith, *supra* note 258, at 288.

²⁷⁷ For example, the aid provision might be included in an omnibus, end-of-session spending bill full of targeted benefits for various kinds of interest groups.

²⁷⁸ See *infra* text accompanying notes 423-27 (presenting a justificatory account of liberal neutrality); cf. Laycock, *supra* note 270 (presenting a different account of Establishment Clause neutrality).

In short, the “endorsement” test is, at best, nested within a disjunctive account of the Establishment Clause that includes reference to nonexpressive factors as well as to the sentence meaning of a governmental action. The expressivist can claim at best that a law violates the Establishment Clause if the law (a) breaches relevant nonexpressive factors such as neutrality, coercion, etc. *or* (b) triggers a linguistic convention for performing the speech-act “Government endorses Religion *R*” or “Government endorses religion in general.”²⁷⁹

Will this kind of disjunctive account prove successful? Note, to begin, that the second, expressive prong of the disjunctive structure risks collapsing into the first. For example, if the factors listed under (a) include neutrality (and not just, say, coercion), then it is difficult to see what further analytic work is done by the endorsement test. Seemingly, the set of “endorsements” is wholly subsumed within the set of governmental actions that are nonneutral, and thus prong (b) seems quite superfluous.²⁸⁰ Let us, however, bracket this issue. (Perhaps the expressivist can say that nonneutral endorsements, as opposed to meaningless acts of nonneutrality, are *particularly* bad; or perhaps the first, nonexpressive factor is more limited in scope than neutrality.) The deeper problem that the expressivist must address is what motivates prong (b). Why does the linguistic meaning of a governmental action figure in the best or most perspicuous description of that action, within the context of the Establishment Clause? What is morally distinctive about “endorsement”?

Interestingly, Justice O’Connor’s answer to this question tracks

²⁷⁹ See Karst, *supra* note 76, at 517 (“Problems both definitional and institutional embarrass the effort to make the endorsement of religion into a generalized legal test governing the whole range of Establishment Clause issues. . . . [These] problems diminish, however, when the case at hand involves an official governmental display of a religious symbol.”). Justice O’Connor herself now concedes as much. See *Capitol Square Review & Advisory Bd. v. Pinette*, 515 U.S. 753, 773–74 (1995) (O’Connor, J., concurring) (arguing that the Establishment Clause cannot be reduced to a single test and apparently proposing that “endorsement” be given a more limited role than in *Lynch v. Donnelly*: “the endorsement inquiry captures the fundamental requirement of the Establishment Clause when courts are called upon to evaluate the constitutionality of religious symbols on public property”); *Board of Educ. v. Grumet*, 512 U.S. 687, 718–21 (1994) (O’Connor, J., concurring) (arguing that the Establishment Clause cannot be reduced to a single test).

²⁸⁰ The claim I am making here is *not* that the property of endorsement is necessarily parasitic upon nonexpressive properties, see *supra* text accompanying notes 271–72 (criticizing this claim), but rather that any governmental action genuinely “endorsing” religion—in a genuine, linguistic way—is likely unconstitutional in virtue of some non-expressive property, specifically nonneutrality, that could attach to meaningless actions as well. It is hard to imagine how a law could possess the linguistic meaning “We endorse Religion *R*” or “We endorse religion in general” and still be neutral.

very closely the expressivist account of the Equal Protection Clause offered by Paul Brest and others, as well as the expressive theory of punishment offered by Jean Hampton. Just as equal protection expressivists such as Brest focus on stigma, and just as Hampton focuses on the reversal of victim "diminishment," so O'Connor argues that endorsement makes second-class citizens of those persons who do not share the endorsed religious beliefs. Thus her oft-quoted concurring analysis in *Lynch*: "Endorsement sends a message to nonadherents that they are outsiders, not full members of the political community, and an accompanying message to adherents that they are insiders, favored members of the political community. Disapproval sends the opposite message."²⁸¹ The idea, once again, is that a governmental message can affect the status and concomitant self-respect of certain persons. Endorsement of religion *R*, or of religion in general, causes status harm to nonbelievers.

This has been the standard defense of the endorsement test, not just by Justice O'Connor, but also by its advocates within the legal academy. Gary Leedes asserts that the Establishment Clause and the endorsement test "prohibit[] the federal and state governments from subverting a citizen's status in the political community because of his or her creed or lack of religious commitment."²⁸² Arnold Loewy argues that the test "is well suited to preventing successful government attempts . . . to impose a 'badge of inferiority' on our religious minorities."²⁸³ Kent Greenawalt suggests that it seeks "to avoid feelings of exclusion and dominance."²⁸⁴ And Kenneth Karst, a leading stigma theorist within the Equal Protection Clause context, has simply extended his theory to include the Establishment Clause:

In the generation of lawyers that has so decisively repudiated Jim Crow, two lessons are clear First, governmental expression has a considerable capacity to alienate outsiders. Second[,] . . . this form of alienation is, in itself, a harm of major proportion. When government sponsors the symbols of religion, the spoils . . . are mainly psychic.²⁸⁵

²⁸¹ 465 U.S. 668, 688 (1984) (O'Connor, J., concurring). My Westlaw search, on April 25, 2000, discovered some 198 law review articles citing this language or a portion thereof.

²⁸² Leedes, *supra* note 76, at 469.

²⁸³ Loewy, *supra* note 76, at 1069.

²⁸⁴ Greenawalt, *supra* note 251, at 374.

²⁸⁵ Karst, *supra* note 76, at 512. Marshall, *supra* note 260, defends endorsement on different grounds, namely the incoherence of any objective approach to the Establishment Clause. See *id.* at 531-33. If this is true, it is unclear what would justify the Court in invalidating endorsements. Why not just declare the clause an inkblot?

But the difficulty with this line of argument, once more, is that the semantics of some governmental utterance are not equivalent to its status (or more broadly its cultural) impact. Government endorses religion when it invokes a particular linguistic convention; it changes a person's status when it changes the prevalent beliefs about his inferiority or equality, or the prevalent treatment of him motivated by beliefs about his inferiority or equality, or the beliefs or treatment that are appropriate pursuant to existing social practices. The two are as different as, in general, law and social norms are. A governmental action with the linguistic meaning constitutive of endorsement is neither necessary nor sufficient to cause status harm to the nonbeliever—to damage his standing as a full member of the political community.

Professor Smith, perhaps the leading scholarly critic of the endorsement test, has quite vigorously pursued this particular line of criticism.²⁸⁶ As for necessity, Smith gives the example of state laws that excluded clergy from serving in state legislatures, laws that were quite prevalent until struck down by the Court in 1978.

These laws plainly affected some persons' political standing on the basis of religion; the exclusionary laws made those persons ineligible for legislative office simply because they had chosen a religious vocation. On the other hand, whether the laws communicated approval or disapproval of religion is debatable Such a law *might* reflect disapproval of religion, implying that ministers are unfit for public office. Conversely, the law might suggest *approval* of religion; it might evince a belief that ministers are too virtuous . . . to be sullied and distracted by mundane political pursuits. Or the law might reflect neither approval nor disapproval of religion, but merely a belief that both religion and politics are better off when kept apart.²⁸⁷

As for sufficiency, Smith points out that "[c]eremonial uses of prayer [by governmental officials] may communicate support or approval for religious beliefs[,] [b]ut . . . no one loses the right to vote, the freedom to speak, or any other [political] right[s]."²⁸⁸ This is probably too strong an assertion. I would think that ceremonial prayer, or some other endorsement without prescriptive meaning (that is, without any effect on legal rights, duties, etc.), could, in the

²⁸⁶ See Smith, *supra* note 258, at 306 ("To be sure, a law diminishing or elevating the political standing of citizens on religious grounds *might* also endorse or disapprove of religion, and vice versa. But those consequences of such a law are practically and analytically distinct.").

²⁸⁷ *Id.* at 306-07 (footnotes omitted).

²⁸⁸ *Id.* at 307.

right circumstances, diminish and alienate nonbelievers.²⁸⁹ But what about an endorsement of religion *R*, uttered by a government under the control of legislators from that religion, when the vast majority of the citizenry are nonbelievers? (Imagine that the Unification Church succeeds in converting a majority of the legislature.) Or, more plausibly, what about an endorsement of religion *R*, or of religion in general, uttered by a government about which citizens are generally disillusioned, and whose officials are generally viewed as corrupt and petty? The expressivist must either insist on counting that statement as a distinctive moral wrong, notwithstanding its cultural irrelevance, or she is no longer an expressivist in the sense that links "expression" to language and linguistic meaning.²⁹⁰

C. *Expressive Theories of Regulation*

An expressive theory of regulation, or at least the outline of such a theory, is presented in scholarship by Professors Sunstein, Pildes, and Anderson, particularly Sunstein's article *Incommensurability and Valuation in Law*,²⁹¹ a joint article by Pildes and Anderson entitled *Slinging Arrows at Democracy*,²⁹² and a joint article by Sunstein and Pildes entitled *Reinventing the Regulatory State*.²⁹³ This theory is less well developed than the expressive theories of punishment and of constitutional law discussed above in Parts II.A and II.B. Anderson, Pildes, and Sunstein are (quite understandably) more concerned to criticize the standard, nonexpressive account of regulation than to develop a specific, expressivist alternative. Nonetheless, their inchoate theory is worth considering here, both because of its potential import, and because the

²⁸⁹ See generally Epstein, *supra* note 117, at 2124-54 (arguing that ceremonial deism makes outsiders of nonbelievers, and thereby violates the Establishment Clause).

²⁹⁰ Here, as in the case of the Equal Protection Clause, the notion of endorsement might play a *doctrinal* role without constituting a morally basic factor. See *supra* text accompanying notes 243-45 (distinguishing between expressive cast of moral theory and expressive cast of doctrine). But again, the existence of a moral case for linguistic meaning to play a doctrinal or, more generally, a legal role is not, yet, genuine expressivism. See *id.*

²⁹¹ Sunstein, *Incommensurability and Valuation*, *supra* note 16.

²⁹² Pildes & Anderson, *Slinging Arrows*, *supra* note 8.

²⁹³ Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8. I say "particularly" because these scholars have written widely, and the theory sketched here is, not surprisingly, evidenced in some of their work beyond these three articles. See, e.g., ANDERSON, *VALUE IN ETHICS AND ECONOMICS*, *supra* note 143, at 190-220; Richard H. Pildes, *The Unintended Cultural Consequences of Public Policy: A Comment on the Symposium*, 89 MICH. L. REV. 936 (1991); Sunstein, *Incommensurability and Kinds of Valuation*, *supra* note 16.

arguments that they offer to support it are, in certain respects, quite different from those developed in the expressivist literature on punishment, the Equal Protection Clause, and the Establishment Clause.

Numerous passages in the articles just mentioned suggest that Anderson, Pildes, and Sunstein are indeed genuine expressivists—that they characterize the moral factors applicable to governmental regulation as including some irreducibly expressive factors. Consider the following statement from *Reinventing the Regulatory State*:

[One] problem with [cost-benefit] approaches is that they necessarily focus on the quantitative or material effects of policies. They cannot take into account what we will call *the expressive dimensions of legal and political choices*. By expressive dimensions—what might be understood as cultural consequences of choice—we mean the values that a particular policy choice, in the specific context in which it is taken, will be generally understood to endorse. Policy choices do not just bring about certain immediate material consequences; they also will be understood, at times, to be important for what they reflect about various value commitments—about which values take priority over others, or how various values are best understood. Both the material consequences and the expressive consequences of policy choices are appropriate concerns for policymakers.²⁹⁴

Pildes and Sunstein continue: “A society might protect endangered species partly because it believes that the protection makes best sense of its self-understanding, by expressing an appropriate valuation of what it means for one species to eliminate another.”²⁹⁵ Similarly, Pildes and Anderson have this to say about the inadequacy of a non-expressive, consequentialist account of the criteria bearing on regulatory choice:

Some public values are “hierarchically” incommensurable with others, meaning that the incomparably higher regard for one value over the other is expressed by refusing certain types of trade-offs between the two. Social choice theory cannot represent the relations of hierarchically incommensurable values because an individual’s attitudes toward them cannot be captured in terms of a single consequentialist preference

²⁹⁴ Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 66. To be sure, the quoted phrase, “what might be understood as cultural consequences of choice,” cuts against an interpretation of Pildes and Sunstein that equates “expressive” with “linguistic,” and in favor of an interpretation that equates it with “cultural.” Nonetheless, the overall quotation here, as well as other references to “expression” in *Reinventing the Regulatory State*, is sufficiently ambiguous on this score that the linguistic-meaning variant is, at the least, a plausible reading of the theory that Sunstein and Pildes mean to present. My criticisms are directed towards this variant and are *not* meant to show that a cultural-impact theory of regulation is similarly unpersuasive.

²⁹⁵ *Id.* at 69.

ranking. Any description of options in terms of their consequences alone, apart from their expressive significance, will exclude some of the concerns individuals have that influence their choices.²⁹⁶

Pildes and Anderson point to “life” and “money” as examples of moral considerations whose connections cannot be captured in a nonexpressive framework. They argue that “the best way to express the hierarchically superior value of life over money is not in a consequentialist, lexical preference ranking of ‘life’ over ‘money,’ but by the principle to express respect for life under all circumstances. On its face, this principle inherently involves expressive concerns”²⁹⁷

More specifically, I suggest, Anderson, Pildes, and Sunstein all seem to endorse the following three propositions.

(1) *Pluralism*

There is a plurality of *values*. As Sunstein puts it: “[H]uman values are plural and diverse. . . . [W]e value things, events, and relationships in ways that are not reducible to some larger and more encompassing value.”²⁹⁸ Similarly, Pildes and Anderson argue for “value pluralism,” which they characterize as “the view that public values—those values at stake in political choice—ought to be understood to be diverse in a particular and profound way.”²⁹⁹

(2) *Expressive constraints*

Public officials, in particular regulators, are constrained to express an *appropriate evaluative view*—an appropriate understanding of what values (or some values) require, and of how values (or some values) relate. Pildes and Sunstein state: “Policy choices do not just bring about certain immediate material consequences; they also will be understood, at times, to be important for what they reflect about various value commitments—about which values take priority over others, or how various values are best understood.”³⁰⁰ And Pildes and Anderson have the following to say about the hierarchical, expressive structuring of values: “When higher values are at stake, *particular kinds* of comparisons with lower values are considered inappropriate, immoral, or unjust—comparisons that would *express* a degradation or depreciation of the higher values.”³⁰¹

²⁹⁶ Pildes & Anderson, *Slinging Arrows*, *supra* note 8, at 2146.

²⁹⁷ *Id.* at 2151.

²⁹⁸ Sunstein, *Incommensurability and Valuation*, *supra* note 16, at 780.

²⁹⁹ Pildes & Anderson, *Slinging Arrows*, *supra* note 8, at 2143.

³⁰⁰ Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 66.

³⁰¹ Pildes & Anderson, *Slinging Arrows*, *supra* note 8, at 2150.

(3) *Incommensurability*

Because the criteria for regulatory choice are, in part, irreducibly expressive, no nonexpressive scale for integrating the plural values bearing upon choice will capture all the relevant considerations. Options are *incommensurable*, relative to such scales (for example, relative to the money scale employed by cost-benefit analysts).³⁰² As Sunstein puts it: "Incommensurability occurs when the relevant goods cannot be aligned along a single metric without doing violence to our considered judgments about how these goods are best characterized."³⁰³ Since a cost-benefit metric, as standardly defined, incorporates only nonexpressive considerations, regulatory options implicating expressive constraints will be incommensurable by that metric:

CBA [cost-benefit analysis] approaches cannot adequately capture all the expressive dimensions of policy choices. . . . CBA deals with the material or quantitative dimensions, not the interpretive and expressive ones. CBA examines alternative end states; it compares, for example, how much it would cost to reach a state in which health was protected to a certain degree against a particular risk. It cannot take account of the meaning of the transition—the values the transition will be socially understood to express—from one end state to another.³⁰⁴

What is distinctive about the views of Anderson, Pildes, and Sunstein, and what (I think) justifies the claim that they are offering a common, expressivist theory of regulation, is the union of these three components: (1) pluralism, (2) expressive constraints, and (3) incommensurability.

So where does the theory go wrong? Note, to begin, that it is a mistake to infer the second component of the theory from the first. To the extent that Anderson, Pildes, and Sunstein mean to claim such an inference, they are mistaken. The plurality of values does not imply that *expressive* criteria govern, or partly govern, regulatory choices implicating plural values. Pluralism, as such, is fully consistent both with expressivism and with the absence of any expressive factors.³⁰⁵

³⁰² See Adler, *Introduction*, *supra* note 17, at 1170, 1177-80 (describing the multiple meanings of "incommensurability," and suggesting that incommensurability in one sense means "the failure of a particular kind of scale . . . to track the comparative worth of options").

³⁰³ Sunstein, *Incommensurability and Valuation*, *supra* note 16, at 796 (emphasis deleted).

³⁰⁴ Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 70.

³⁰⁵ The fact and implications of pluralism are discussed throughout the philosophical literature on incommensurability, *see* Adler, *Cost-Benefit Analysis*, *supra* note 17, at 1383 n.47 (citing literature), but particularly in MICHAEL STOCKER, *PLURAL AND CONFLICTING VALUES* (1990) and JOHN KEKES, *THE MORALITY OF PLURALISM* (1993).

To see this point, consider the following. First, a "value" might be defined in a fairly standard way as a nonexpressive criterion for ranking the outcomes that actions produce—a criterion such that meaningless or linguistically identical actions can fare differently with respect to the criterion.³⁰⁶ "Physical pleasure" or "human longevity" or "the preservation of endangered species S_i " constitute "values" in this fairly standard sense: two meaningless or linguistically identical actions, by government, can produce world-states that differ with respect to the aggregate amount of physical pleasure, the extent of premature death, or the number of members of species S_i . It is very hard to see why the sheer plurality of values, thus defined, should give rise to expressive constraints. *Universe₁* is a monistic universe; only one nonexpressive criterion for ranking outcomes obtains. *Universe₂* is a pluralistic universe; several nonexpressive criteria for ranking outcomes obtain. Why should the shift from *Universe₁* to *Universe₂* trigger the emergence of expressive constraints?³⁰⁷

³⁰⁶ See Kornhauser, *supra* note 116, at 1604 (defining "value" in roughly this sense). One might quibble with my definition and say that values in a standard sense are possibly (not necessarily) nonexpressive, that outcome criteria may or may not have an expressive cast—for example, the value of "government's saying ___"—but this qualification does not materially change my argument. Assume a definition of value such that values may or may not be expressive. Once more, the pluralism/monism debate turns out to be orthogonal to the debate about expressivism. A monistic universe can be expressive, or not, depending on whether its one value is expressive; and the same is true of a pluralistic universe, depending on whether one or more of its several values are expressive.

³⁰⁷ The expressivist might respond that, absent expressive factors, there is simply no way to make trade-offs between plural values (i.e., plural, nonexpressive criteria for ranking outcomes). Let us put the argument this way: "If expressive constraints did not exist, then Global Incomparability would be true. Outcome O^* could not be ranked as all-things-considered better than outcome O unless O^* was better than O with respect to each and every value $V_1, \dots, V_n$. But Global Incomparability is massively counterintuitive. Therefore, expressive constraints exist." This is the best argument that I can imagine for inferring expressivism from pluralism. However, the argument is wrong. Global Incomparability is massively counterintuitive but it is untrue that, apart from expressive considerations, there is no way to make warranted trade-offs between plural values. See, e.g., THOMAS HURKA, *PERFECTIONISM* 84-98 (1993) (arguing against Global Incomparability); LARRY S. TEMKIN, *INEQUALITY* 141-56 (1993) (same); Adler, *Cost-Benefit Analysis*, *supra* note 17, at 1402-04 (same); Ruth Chang, *Introduction*, in *INCOMMENSURABILITY*, *supra* note 16, at 14-17 (same); James Griffin, *Incommensurability: What's the Problem?*, in *INCOMMENSURABILITY*, *supra* note 16, at 38-40 (same); Donald Regan, *Value, Comparability, and Choice*, in *INCOMMENSURABILITY*, *supra* note 16, at 134-39 (same).

Imagine that O^* saves 10,000 lives relative to O , but that O increases the number of habitats of an endangered species, relative to O^* . Surely we can say that O^* is, all things considered, better than O , quite apart from what O "says" about life. Imagine that the O - O^* choice is subsumed by no linguistic conventions; it has no sentence

Second, a "value" might be defined in a less standard way, such that any value is identical to or includes an expressive constraint. For example, a value V_i might be the following: a hybrid of (1) a nonexpressive criterion for ranking outcomes and (2) an expressive obligation, to the effect that actors communicate an appropriate view about what V_i requires. But if this is what values involve, then once more, it is not clear why pluralism has anything to do with expressivism. If monistic *Universe₁* contains a single, partly expressive value—say, the value of life, understood to incorporate a requirement that actors communicate respect for life—then expressive constraints *do* obtain in *Universe₁*, as they also do in pluralistic *Universe₂*, which includes both the value of life as well as other values.

In short, if values are (standardly) defined as nonexpressive criteria for ranking outcomes, then the pluralism/monism debate is orthogonal to the expressivism debate, in that—regardless of whether values are plural or single—it remains a further and contingent question whether expressive constraints obtain. If values are (nonstandardly) defined as equaling or incorporating expressive constraints, then expressivism is true in both pluralistic and monistic worlds. In either event, the fact of pluralism—and I would agree with Anderson, Pildes, and Sunstein that values are, in fact, plural³⁰⁸—does not advance the case for expressivism.

A parallel point can be made about the incommensurability component of the Anderson, Pildes, and Sunstein theory. Incommensurability is surely an *upshot* of expressivism. If the moral criteria for regulatory choice are, in part, irreducibly expressive, then, indeed, no nonexpressive scale for ranking governmental actions will capture all

meaning, apart from its prescriptive meaning. Surely, still, O^* is better than O . Nor should we be tempted to say that O necessarily symbolizes or expresses contempt or disregard for human life—that its meaning is more than a contingent matter of linguistic conventions—for the only sense in which it necessarily expresses contempt or disregard is that it purchases a meager gain in environmental preservation at the expense of a massive loss in life. It necessarily possesses this "meaning" in the Gricean, signaling sense, only because meager gains to the first value are less important than massive losses to the second value quite apart from the linguistic meaning of that choice.

It is true or at least quite plausible that in a universe of plural, nonexpressive, outcome criteria, *some* choices would be unguided. Call this "Local Incomparability." See generally Chang, *supra*, at 13-27 (summarizing arguments for incomparability, particularly Local Incomparability); Kornhauser, *supra* note 116, at 1599-1622 (discussing the connection between pluralism and incomparability). But this is not particularly unsettling or weird and does not get us from pluralism to expressivism.

³⁰⁸ See, e.g., KEKES, *supra* note 305 (arguing for plurality of values); STOCKER, *supra* note 305 (same); Chang, *supra* note 307, at 14 (citing sources on this issue).

the relevant moral considerations. Further, incommensurability may well be an upshot of pluralism. The existence of plural values may always lead to incomparability—to options that are neither better, nor worse, nor precisely equally good with respect to the totality of values—and no scale of any kind can represent the incomparability of options.³⁰⁹ Anderson, Pildes, and Sunstein are correct in drawing the connection *from* expressivism and pluralism *to* incommensurability. But the demonstration of that connection does nothing to advance the case for expressivism. We still need to know why, to begin with, governmental actors are subject to expressive obligations—why they are constrained to communicate an appropriate understanding of what values (or some values) require, and how values (or some values) relate.

I should stress that Anderson, Pildes, and Sunstein may not intend to make the claim that expressivism follows from pluralism, incommensurability, or a combination thereof. Whether they do is unclear to me. In any event, the claim is spurious.

What else, if anything, do Anderson, Pildes, and Sunstein have to say in favor of expressive constraints? First, there is some apparent reference to Anderson's expressive theory of individual action, presented in *Value in Ethics and Economics*.³¹⁰ But as I argued in Part I, the inference from an expressive theory of individual action to an expressive theory of governmental action is a false one.³¹¹ The central idea in Anderson's expressive theory of individual action is that individuals have relationships with particular others, which relationships persons act to express. "Practical reason demands that one's actions adequately express one's rational attitudes toward the people and things one cares about."³¹² Yet government has no special relationship to particular persons, excepting (perhaps) its special relationship to citizens as against noncitizens³¹³—and even this is different from the

³⁰⁹ More precisely, pluralism plausibly leads to Local Incomparability. See *supra* note 307.

³¹⁰ See ANDERSON, *VALUE IN ETHICS AND ECONOMICS*, *supra* note 143; Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 69 (discussing individual behavior); Sunstein, *Incommensurability and Valuation*, *supra* note 16, at 822-23 (same); *id.* at 821 n.141 (claiming to be "indebted" to *Value in Ethics and Economics*).

³¹¹ See *supra* Part I.E.

³¹² ANDERSON, *VALUE IN ETHICS AND ECONOMICS*, *supra* note 143, at 18.

³¹³ Cf. DWORKIN, *supra* note 106, at 195-215 (describing the attitude of "equal concern" that fellow citizens owe to each other and concluding that "nothing in this argument suggests that the citizens of a nation state, or even a smaller political community, either do or should feel for one another any emotion that can usefully be called love").

kinds of familial or affectionate relationships that are central to *Value in Ethics and Economics*, since government's relationship with citizen *P* is not a relationship with him, come what may, but with him under the description "citizen," which evaporates once that description no longer obtains.

A possible response, I suppose, is that governmental officials might be obliged (1) to adopt one or another attitude towards some person, and (2) to act to express that attitude, even though those officials have no special relationship to that person. For example, governmental officials are plausibly obliged to adopt an attitude of *respect* towards all persons, citizens and noncitizens alike. Still, it remains the case that the set of attitudes appropriately adopted by governmental officials is much smaller than the set appropriately adopted by individuals. Consider the following exemplary list of "basic evaluative attitudes" presented at one point in *Value in Ethics and Economics*: "love, respect, consideration, affection, honor, and so forth."³¹⁴ I would think it improper for governmental officials to be motivated, in that capacity, by love or affection, and at least problematic for them to be motivated by consideration and honor.

If there is only a small range of attitudes that governmental officials appropriately adopt, then the theory of rationality propounded in *Value in Ethics and Economics*—rational action is action that expresses the attitudes appropriately held by the actor—is implausible as a theory of rational (or moral) action by government. "Express your appropriate attitudes" (or that injunction together with the injunction, "Refrain from expressing inappropriate attitudes") is surely not the *only* moral injunction binding upon governmental officials.³¹⁵ In-

³¹⁴ ANDERSON, *VALUE IN ETHICS AND ECONOMICS*, *supra* note 143, at 20.

³¹⁵ Lest I be accused of adopting an idiosyncratic and unfair reading of *Value in Ethics and Economics*, I would refer the reader to John Broome's review of the book. See John Broome, *Review*, 9 *RATIO (NEW SERIES)* 90 (1996). Broome interprets *Value in Ethics and Economics* as proposing that "the aim of actions is to *express* valuations [with] 'expression' . . . to be taken literally." *Id.* at 92. He then goes on to object:

I cannot believe this theory. When I value something, what does my value demand of my actions? Sometimes, no doubt, it demands expression. . . . But much more often values demand to be *achieved*. When I hand a child a life-jacket as she gets on my boat, I dare say that expresses my concern for her safety. But I do not do it for that reason. I do it because it makes her safer; it helps to achieve the value of safety. . . . If I were alone in the world, with no one to express my values to, I would still need to act rationally so as to achieve my values. Anderson has picked out one important but minor role of actions and treated it as the whole.

Id. I happen to think that Broome's point constitutes a fatal criticism of the theory of individual rationality proposed in *Value in Ethics and Economics*, but whether or not that

deed, a close reading of the Anderson, Pildes, and Sunstein articles on regulation suggests that the authors do not really intend to apply the *Value in Ethics and Economics* theory to governmental action, because the articles concede that expressively identical or meaningless governmental actions can differ in goodness or rightness.³¹⁶ We might adopt a weaker expressive theory of governmental action—namely, governmental actors must (1) comply with applicable nonexpressive requirements and (2) act to express the rational attitudes (e.g., respect) that they do appropriately hold—but this is not the theory advanced by *Value in Ethics and Economics*,³¹⁷ and not a theory that the Pildes, Anderson, and Sunstein articles discussed here attempt to defend.³¹⁸

A second, brief argument for expressivism comes with an invoca-

is true, I think it constitutes conclusive grounds against extending *Value in Ethics and Economics* to governmental action: government is obliged to achieve values, not just to express them. At the very least, this is true if “express” is used in the way I am using it in this Article, to refer to the communication of linguistic meaning.

There is some possibility that *Value in Ethics and Economics* is using “express” and “expressive” in a different, nonlinguistic sense. If so, the “expressive” (nonlinguistic) theory presented by *Value in Ethics and Economics* presents no challenge whatsoever to my claim that “expressive” (linguistic) theories of regulation and, more generally, of law are unpersuasive.

³¹⁶ See, e.g., Pildes & Anderson, *Slingshot Arrows*, *supra* note 8, at 2153 (“In evaluating options, we look *not only* at the consequences of our choices—the actual combinations of safety and money achieved by them—but at the expressive significance of making these choices in the context at hand.” (emphasis added)); *id.* at 2145–46 (providing an exemplary list of values, which suggests a conceptualization of values as equaling or including nonexpressive criteria for ranking outcomes or actions: “[m]ost interesting individual or collective decisions implicate a wide diversity of values [such as] . . . respect[ing] rights, fulfill[ing] duties, satisfy[ing] needs, carry[ing] out promises, improv[ing] welfare, and follow[ing] through on the special projects and commitments that we have collectively decided to undertake”); Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 66 (stating that “[b]oth the material consequences and the expressive consequences of policy choices are appropriate concerns for policymakers”); *id.* at 74–75, 127 (suggesting that policy assessment might be performed in two stages, with quantitative cost-benefit analysis undertaken at the first stage and expressive issues considered at the second); Sunstein, *Incommensurability and Valuation*, *supra* note 16, at 824 (“I do not claim that the expressive effects of law . . . are decisive or that they cannot be countered by a demonstration of more conventional bad consequences.”).

³¹⁷ See Broome, *supra* note 315, at 92.

³¹⁸ Pildes has recently disclaimed interest in the actual attitudes or other mental states held by governmental officials. See *supra* note 91 and accompanying text. The Pildes and Anderson article under discussion here itself does so, at least in the context of statutory interpretation. See Pildes & Anderson, *Slingshot Arrows*, *supra* note 8, at 2207–08 (“Apart from the ways in which they are mediated by [Congress’s] institutional practices [e.g., the practice of issuing a committee report to accompany a statute], individual subjective purposes do not have any particular role in the construction of collective purposes and public meanings expressed in statutes . . .”).

tion of the value of "integrity." Pildes and Sunstein suggest that:

[T]here is a . . . ground for endorsing the expressive function of law [which] is not about social effects [i.e., instrumental effects on social norms] in the same sense. To understand this idea, it is helpful to start with the personal interest in integrity. . . . [W]e might say that individual behavior is not concerned solely with states of affairs, and that if it were, we would have a hard time making sense of important aspects of our lives. Personal integrity, commitment, and the narrative continuity of a life matter enormously as well.³¹⁹

Do Pildes and Sunstein really intend to say that, if one regulatory choice is better justified than another on nonexpressive grounds, government might nonetheless be all-things-considered justified in choosing the latter because of considerations of *governmental* integrity, commitment, and narrative continuity? Perhaps they do; after all, Ronald Dworkin has famously argued for integrity as a virtue of legal systems.³²⁰ But Dworkin's "integrity" is nonexpressive—it is about legal decisions truly cohering with each other, not about decisions expressing the importance of coherence³²¹—and Pildes and Sunstein do not begin to show that, *pace* Dworkin, governmental integrity, commitment, and narrative continuity are best understood symbolically.

A third attempt to demonstrate the existence of expressive constraints comes with the invocation of deontological rules. Pildes and Sunstein, after discussing the virtues of personal integrity, commitment, and the narrative continuity of a life, proceed to observe that "someone might refuse to kill an innocent person at the request of a terrorist, even if the consequence of the refusal is that many more people will be killed. Our responses to this case are not adequately captured in purely consequentialist terms."³²² This is absolutely cor-

³¹⁹ Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 69; *see also* Sunstein, *Incommensurability and Valuation*, *supra* note 16, at 822-23 (similarly arguing for connection between expressivism and personal integrity).

³²⁰ *See, e.g.*, DWORKIN, *supra* note 106, at 175-275.

³²¹ Although Dworkin does refer to the "expressive value" of integrity, *see id.* at 189, 190, I take the accomplishment of this value to derive from (indeed, to reduce to) the satisfaction of the nonexpressive requirements of fit, coherence, and so on that Dworkin groups under the rubric of integrity. For Dworkin, the internal coherence of a community's laws expresses its identity as a community of principle. This is quite different from saying that integrity demands certain *statements* by lawmakers. *See, e.g., id.* at 190 ("[T]he expressive value [of integrity] is confirmed when people in good faith try to treat one another in a way appropriate to common membership in a community governed by political integrity . . . even when they disagree about exactly what integrity requires in particular circumstances.").

³²² Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 69; *see also* Sunstein, *Incommensurability and Valuation*, *supra* note 16, at 822-23 (similarly invoking

rect. The rule against killing is *the* paradigmatic example of a non-consequentialist moral constraint. That constraint is both deeply plausible—indeed, we ought not kill—and nonconsequentialist in form, since we are prohibited from killing one person even to prevent five killings.³²³ But unless deontological constraints such as the no-killing constraint are, at bottom, rules against *saying* something, it is a non sequitur to leap from the genuine existence of such constraints to the conclusion that expressive factors also obtain.

In fact, there are strong arguments against an expressivist account of deontological constraints. For example, I may not intentionally cause your death, nor may the government intentionally cause your death, even to prevent five future killings, and even if the killing of you has zero symbolic content.³²⁴ The no-killing constraint is a constraint upon doing harm, or upon intentionally causing harm, not upon expressing disrespect for life. This issue is pursued at greater length in Part III, where I argue in a general way that the consequentialism/deontology debate is orthogonal to the debate about expressivism.³²⁵

Finally, and most extensively, Pildes and Anderson argue that the relation of “higher” values (such as the value of life) to “lower” values (such as the value of money) is captured by an expressive requirement that actors must never communicate disrespect for the higher value (e.g., by comparing it to the lower value in the wrong sort of way). They defend this claim through a detailed examination of the Ford Pinto case:

After Ford recognized that the peculiarly designed location of the Pinto’s gas tank subjected drivers and passengers to substantially greater than usual risks of burn injuries or death in low-speed, rear-end collisions, Ford . . . [conducted] a cost-benefit analysis in which a \$200,000 value was assigned to a human life Ford then calculated the total costs of repairs for all vehicles at \$11 per vehicle, found these repair costs to be substantially greater than the benefits of avoiding deaths and burn injuries, and declined to make any repairs.³²⁶

deontological rules).

³²³ See KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 70-105 (discussing the no-killing prohibition); *id.* at 70-152 (discussing deontological constraints).

³²⁴ This example is designed to fall under both of the leading accounts of the no-killing prohibition, namely as the central case of a constraint against *doing* harm (as opposed to merely allowing harm) and, alternatively, as the central case of a constraint against *intending* harm (as opposed to merely doing or allowing harm nonpurposively). See *id.* at 100-05 (discussing and contrasting these two accounts).

³²⁵ See *infra* text accompanying notes 393-415.

³²⁶ Pildes & Anderson, *Slinging Arrows*, *supra* note 8, at 2150-51.

Pildes and Anderson conclude that the public outrage against Ford is best understood "to reflect an [expressive] set of judgments, not about whether trade-offs between safety and expense should ever be made, but about what it meant to make that particular trade-off in the particular context in which it [was] made."³²⁷

This is a false dichotomy. There are numerous nonexpressive accounts of the public's justified moral outrage about the Pinto affair, other than the implausible and incorrect view that life takes lexical priority over money. Ford's valuation of life, \$200,000, was at least ten times too low,³²⁸ so that the decision not to repair might well have been wrong even in straight utilitarian terms. Further, because Ford was marketing the car to consumers unaware of the defect, non-utilitarian (and nonexpressive) considerations of desert and responsibility might have required Ford to repair it even if doing so harmed Ford stockholders more than it helped auto consumers.³²⁹ Finally, if consumers were significantly poorer than stockholders, distributive considerations might have come into play.³³⁰

Indeed, Pildes and Anderson fail to show that Ford's decision had expressive content—that it implicated linguistic, sentence-meaning conventions or had a real, Gricean speaker's meaning. They claim that "given the background of social and legal understandings against which Ford executives had acted, Ford's particular trade-off *expressed contempt* for human life. . . . To market deliberately or refuse to recall a dangerously defective good expresses contempt for human life . . . [particularly] when the defect is concealed from consumers."³³¹ The reference to contempt, here, just makes the account needlessly complicated. We go from background rules constraining certain behavior, to an inference that Ford's breach of those rules "meant" contempt, to outrage at the contempt. Why not say that Ford acted *wrongfully* by breaching the rules—by deliberately marketing a defective

³²⁷ *Id.* at 2151.

³²⁸ See generally W. KIP VISCUSI, *FATAL TRADEOFFS: PUBLIC AND PRIVATE RESPONSIBILITIES FOR RISK* 51-74 (1992) (surveying the large body of empirical literature valuing death and injury risks in dollars, and concluding that a reasonable dollar value of life is in the range of \$3 to \$7 million).

³²⁹ See, e.g., JULES COLEMAN, *RISKS AND WRONGS* 303-28, 324 (1992) (describing and defending conception of corrective justice such that "the duty of wrongdoers . . . is to repair the wrongful losses for which they are responsible"); SHER, *supra* note 191 (discussing the nonutilitarian factor of desert).

³³⁰ See, e.g., SCHEFFLER, *supra* note 123, at 26-40 (sketching a nonutilitarian variant of consequentialism that incorporates distributive criteria for ranking states of affairs).

³³¹ Pildes & Anderson, *Slingshot Arrows*, *supra* note 8, at 2151.

product whose defects were concealed—and that the outrage was a response to this wrongdoing?

This suggests a further critique of the kind of expressive constraints that Anderson, Pildes, and Sunstein propose. Not only do they fail to present a persuasive argument for the proposed constraints; there is also reason to think that, by virtue of the particular form the constraints take, such an argument would be very difficult to provide. Remember: government is purportedly constrained to express an appropriate view as to “which values take priority over others, or how various values are best understood.”³³² I will interpret this, perhaps unfairly, as an expressive constraint governing the *descriptive* content of governmental choices. If cognitivism about morality is true—and I understand Pildes, Anderson and Sunstein to be, at bottom, cognitivists³³³—then statements about the relative priority of values, and about the best understanding of values, *will* be descriptive statements.

So government is purportedly constrained to utter an appropriate description of public values. Further, I take it, this expressive constraint is not subsumed by the generic constraint against lying nor by the generic requirement that government justify its decisions.³³⁴ Although a moral theory that prohibits “lying” or obliges legal officials to “articulate a public justification” for their actions *is* indeed a genuine expressive theory—the property of being a lie or being a justification are expressive properties, since only linguistically meaningful actions can possess these properties—Anderson, Pildes, and Sunstein seem to be proposing something considerably more robust.

Thus, there will be some cases in which a regulatory decision is fully justified on nonexpressive grounds, is accompanied by an utterance correctly describing those grounds, is not accompanied by any untruths, and yet *still* expresses an inappropriate valuation. But how could this be? What space is left for expressive criteria prohibiting or requiring descriptive utterances? Either (1) government is prohibited from making certain true descriptive statements or (2) government is required to make certain true statements that are not demanded by the generic justification requirement. But (1) is deeply unappealing. Surely Anderson, Pildes, and Sunstein do not mean to prevent gov-

³³² Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 66.

³³³ See *supra* text accompanying notes 67-68 (defending the claim that legal expressivists are best seen as cognitivists, not metaethical expressivists).

³³⁴ See *infra* text accompanying notes 404-11, 432-39 (discussing these requirements).

ernment from stating the truth about what values require or about how values relate. So the Anderson/Pildes/Sunstein constraint seems to be a super-justification requirement, along the lines of (2): Government must utter some descriptive statement beyond the statement identifying the nonexpressive criteria that bear upon and justify its decision. But what would be the content of that additional statement? Descriptive statements which do not serve to justify government's decisions and are not required by the nonexpressive criteria themselves (for example, because the statements enhance overall well-being or equality) would seem to be utterly inert and superfluous.

Let me close by emphasizing, once more, that Anderson, Pildes, and Sunstein do not present a fully developed expressive theory of regulation. Perhaps, when fully developed, it could withstand the criticisms I have here adduced.³³⁵ But it is in any event useful to delineate the serious difficulties that such a theory would need to confront. My discussion is intended to do that. Notably, Sunstein himself seems now to be persuaded that the difficulties are insurmountable. In his article *On the Expressive Function of Law*, which was published subsequent to *Incommensurability and Valuation* and *Reinventing the Regulatory State*, he concedes that "at least for purposes of law, any support for 'statements' [by government] should be rooted not simply in the intrinsic value of the statement, but also in plausible judgments

³³⁵ I should note two themes that have generally emerged in Pildes' work on constitutional law, and that he might perhaps want to deploy in order to bolster an expressive theory of regulation. These are the notions of "expressive harm" and of "exclusionary reason." See Pildes, *Why Rights Are Not Trumps*, *supra* note 4, at 733-60. As for the first: I have already argued, in my discussion of the Equal Protection Clause and the Establishment Clause, that there is only a contingent connection between the stigmatic content of governmental utterances and the occurrence of harm to a person's self-respect (mediated by a change of her status); and I argue more generally below that moral theory, insofar as it is focused on well-being and harm, is not expressive. See *supra* text accompanying notes 235-42, 286-90; *infra* text accompanying notes 337-75. As for exclusionary reasons: it is not completely clear to me whether Pildes intends his "exclusionary reason" account as an argument in favor of an expressive theory of constitutional law or as a separate account. In any event, I do not see how the fact that governmental officials are constitutionally excluded from acting on certain reasons (if indeed they are) supports expressivism in the linguistic sense under consideration here. To say that a particular reason is excluded for a particular person is either to make a *motivational* claim (that the person ought not act on the reason, or give it weight in her deliberations, or something like that) or a *justificatory* claim (that the reason has no weight in justifying the person's action). It does not follow (without more) that the person is required to engage in, or to refrain from engaging in, a particular act of linguistic communication. See Heidi Hurd, *Challenging Authority*, 100 YALE L.J. 1611, 1618-44 (1991) (discussing Raz's concept of an exclusionary reason); Michael S. Moore, *The Works of Joseph Raz: Authority, Law, and Razian Reasons*, 62 S. CAL. L. REV. 827, 849-59 (1989) (same).

about its effect on social norms and hence in 'on balance' judgments about its consequences."³³⁶ For the reasons I have here presented, I think that Sunstein's current views, not his former ones, are correct.

III. WHY EXPRESSIVE THEORIES ARE UNPERSUASIVE: A GENERAL ARGUMENT

Part II described and criticized the genuine expressive theories of law that possess some significant degree of scholarly development and salience: the diverse attempts to show that linguistic meaning is a morally irreducible component of punishment; expressivism about constitutional law, specifically about the Equal Protection Clause and the Establishment Clause; and the expressive theory of regulation developed, at least in outline form, by Professors Anderson, Pildes, and Sunstein. But even if my criticisms of all these particular theories are cogent, expressivism might remain a viable option. The criticisms might be weaker, or just inapplicable, with respect to an expressive theory governing some other area of law or possessing some other justificatory structure.

This Part tries to undercut that possibility. I will contend that a genuinely expressive account, whatever its scope or structure, is morally implausible. The various critical themes running through my discussions of punishment, constitutional law, and regulation will reappear here, in a more general and systematic way, and will be joined by new arguments. I take the following approach: I construct a progressively refined and sophisticated picture of the moral criteria governing (or plausibly governing) legal decisions, starting from the simple core of preference-utilitarianism and then adding new moral considerations or refining existing ones. In this process, I attend carefully to whether, at any point, the developing theory becomes genuinely expressive. The answer, I claim, is that it never does.

* *

Preference-utilitarianism is the paradigm of a moral theory that is *not* genuinely expressive. Expressivists standardly criticize preference-

³³⁶ Sunstein, *On the Expressive Function of Law*, *supra* note 20, at 2045; *see also* Sunstein, *Social Norms and Social Roles*, *supra* note 21, at 965 ("Laws with expressive justifications are most plausibly defended on the ground that they will in fact affect social norms and move them in appropriate directions.").

utilitarianism; indeed, they standardly take it as their main foil.³³⁷ *Utilitarianism* is the view that all agents (including governmental agents) are required to maximize overall well-being. *Preference-utilitarianism* consists of that view, plus the specific claim that well-being consists in the satisfaction of preferences.³³⁸ The preference-utilitarian compares two possible actions, A^* and A , of a governmental official, by asking the following: Does the world-state produced by A^* (call it O^*) better satisfy preferences, overall, than the world-state produced by A (call it O)? There is some question within preference-utilitarianism as to how one conceptualizes and measures the overall satisfaction of preferences when the preferences of different persons conflict—when one person prefers O^* to O and another prefers O to O^* .³³⁹ Whatever the answer to that question, however, it seems clear that a preference-utilitarian view does not involve “expressive factors,” as I defined them in Part I, and is therefore not a genuine expressive theory.

The sole moral factors within a preference-utilitarian theory are the preferences of individual persons. Whether A^* is better than A is solely a function of what P_1 prefers, what P_2 prefers, and so on. Yet P_1 's preference (and P_2 's preference, and P_3 's preference, etc.) is not an expressive factor because its satisfaction or frustration does not necessarily depend (even partially) upon what A^* and A mean. A preference can fix upon *any* feature of world-states; it need not fix exclusively, or even partly, upon linguistic features.³⁴⁰ For example, P_1

³³⁷ See *supra* note 130.

³³⁸ For a discussion of utilitarianism, and specifically preference-utilitarianism, see the sources cited *supra* note 134. For more technical discussions of the nature of a preference, see PREFERENCES (Christoph Fehig & Ulla Wessels eds., 1998); ROBERT AUDI, ACTION, INTENTION, AND REASON 35-55 (1993); S.L. HURLEY, NATURAL REASONS 55-83 (1989); Richard J. Arneson, *Liberalism, Distributive Subjectivism, and Equal Opportunity for Welfare*, 19 PHIL. & PUB. AFF. 158, 161-64 (1990); and Arthur Ripstein, *Preference*, in VALUE, WELFARE AND MORALITY 93 (R.G. Frey & Christopher W. Morris eds., 1993).

³³⁹ See generally INTERPERSONAL COMPARISONS OF WELL-BEING (Jon Elster & John E. Roemer eds., 1991) (providing a comprehensive discussion of methods for comparing well-being of different persons, with a significant focus on the problem of comparison given a preference-based view of well-being); Daniel Hausman, *The Impossibility of Interpersonal Utility Comparisons*, 104 MIND 473 (1995) (summarizing and criticizing proposals for comparing preference-satisfaction of different persons).

³⁴⁰ See JAEGWON KIM, PHILOSOPHY OF MIND 13-14 (1996) (describing “desires,” i.e., preferences, as a basic kind of “propositional attitude” and explaining that “[p]ropositions are said to constitute the ‘content’ of the propositional attitudes”). Although other, less basic kinds of propositional attitudes may be restricted in their possible content, it seems clear that preferences are not. In particular it seems quite clear that preferences are not restricted in their possible content to linguistic propositions.

might prefer simply that his basic biological needs be satisfied. Clearly, the "need-satisfying" property of a governmental action is distinct from its linguistic properties. A^* and A can fare differently with respect to the satisfaction of P_i 's needs, even though both A^* and A are meaningless.

Let us now begin to enrich our moral theory. The view that well-being consists in the satisfaction of preferences is simply one position (and a highly controversial one at that) in the space of possibilities. Views of well-being can be divided into the following categories: (a) preference-based views; (b) hedonic views; (c) objective-list views; and (d) mixed views.³⁴¹ Hedonic views postulate that well-being consists in the realization of one or more types of pleasurable mental states or the avoidance of one or more types of painful mental states.³⁴² Objective-list views postulate that well-being consists in the realization of objective goods such as friendship, intellectual or aesthetic accomplishment, recreation, religion, or participation in communal life.³⁴³ Mixed views combine elements of the various pure views; for example, one might plausibly hold that a person is better off with some world-state if and only if she better realizes the balance of objective goods in that state *and* prefers (or comes to prefer) that state.

Let us hold fixed, for now, the utilitarian cast of our moral theory:

³⁴¹ For comprehensive, philosophical discussions of the nature of well-being, see JAMES GRIFFIN, *WELL-BEING: ITS MEANING, MEASUREMENT AND MORAL IMPORTANCE* 7-72 (1986); L.W. SUMNER, *WELFARE, HAPPINESS AND ETHICS* 45-137 (1996). The categorization I adopt here is close but not identical to the categorizations adopted by Sumner or Griffin, which in turn are close but not identical to each other. See also Mozaffar Qizilbash, *The Concept of Well-Being*, 14 *ECON. & PHIL.* 51 (1988) (providing a recent overview of literature on well-being).

³⁴² See GRIFFIN, *supra* note 341, at 7-10 (discussing hedonism as a variant of a "mental-state" view of well-being); SUMNER, *supra* note 341, at 81-112 (discussing hedonism).

³⁴³ Specifically, these are goods whose realization does not entail the satisfaction of a preference or the realization of a pleasurable mental state. For a defense of the need for this qualification, see Matthew D. Adler & Eric Posner, *Rethinking Cost-Benefit Analysis*, 109 *YALE L.J.* 165, 201 (1999):

[W]e must distinguish between objective goods that entail pro-attitudes, and objective goods that do not. 'Recreation' or 'play' is an objective value that, presumably, entails a pro-attitude. I'm not truly playing a game if I'd prefer not to be. A theory of well-being predicated solely on these kinds of [preference]-entailing goods is [not a genuine objective-list theory].

A theory predicated on preference- or pleasure-entailing goods would, on the categorization adopted here, count as a mixed theory. For general discussions of objective-list theories, see SUMNER, *supra* note 341, at 45-80, and GRIFFIN, *supra* note 341, at 40-72. See also JOHN FINNIS, *NATURAL LAW AND NATURAL RIGHTS* 85-90 (1980) (proposing a particular mix of objective goods); GEORGE SHER, *BEYOND NEUTRALITY: PERFECTIONISM AND POLITICS* 199-201 (1997) (same).

government is still required to maximize overall well-being. If the incorporated view of well-being is changed from a preference-based view to a hedonic, objective-list, or mixed view, does the moral theory become expressive?

Clearly, it does not become expressive just by virtue of adding hedonic elements to the theory. The pleasure-related features of a governmental action, like the preference-related features, are not linguistic features. Pleasure, within the hedonic account, means a distinctive positive feeling that is intrinsic to a mental occurrence and that obtains even if the occurrence is dispreferred or lacks objective value.³⁴⁴

The paradigm instances are the pleasures caused by stimuli such as scratching an itch, being massaged, taking a hot bath, quenching a thirst, using a recreational drug, urinating, defecating, and sexual arousal and orgasm. What these sensations have in common, in virtue of which we distinguish them from physical pain, is just the fact that they feel good. When asked to characterize the peculiar feeling tone of sensory pleasure (or pain) we find, like Bentham, that we have little to say. You either recognize what the intense rush of sexual release has in common with the warm glow induced by a backrub, and what differentiates them both from backache or neuralgia, or you do not.³⁴⁵

A hedonic-utilitarian government, choosing between actions A and A^* , would have some basic list of hedonic factors ($F_1 \dots F_n$) by which to evaluate the world-states (O and O^*) that the actions would produce. But these factors will refer to different types of feelings or sensations that persons have in O and O^* , not to the linguistic meaning of A and A^* . The closest we could bring a hedonic theory to genuine expressivism would be to define some hedonic factor, F_p , as the "experience of government saying [some particular thing]." Yet even this highly rigged and artificial variant of hedonic utilitarianism would not be genuinely expressive, because two meaningless actions might produce different *experiences of linguistic meaning* among the population, depending upon their beliefs about which linguistic conventions exist, their degree of attentiveness to governmental actions, and so on.³⁴⁶

³⁴⁴ "Pleasure" is sometimes understood to entail preference, but pleasure in the sense that permits a distinction between preference-based and hedonic views of welfare—pleasure, meaning a positive *sensation*—does not have this entailment. See generally SUMNER, *supra* note 341, at 98-112 (discussing the possible connections and distinctions between "pleasure" and "preference"). Similarly, "pain" within the hedonic account means a distinctive *negative* feeling that is intrinsic to a mental occurrence and that obtains even if the occurrence is preferred or possesses objective value.

³⁴⁵ *Id.* at 106.

³⁴⁶ For the classic discussion (and criticism) of the way in which hedonic views make well-being depend upon the experiences of welfare subjects, see ROBERT

Objective-list theories of well-being, or mixed theories incorporating objective goods, seem to be a much more promising route for the expressivist. Objective goods, or at least some of them, are plausibly *cultural*. They plausibly depend, in part, upon social norms, social forms, or some other such social practices. And there seems to be a much stronger connection between the linguistic meaning of a government's utterances and social norms, forms, or practices, than between the linguistic meaning of a government's utterances and individual preference-satisfaction or hedonic experience. Perhaps the best example of a welfare theory of this kind, incorporating both objective goods and a cultural or social understanding of those goods, is the theory articulated by the well-known moral and legal philosopher, Joseph Raz. Raz claims that "[i]n large measure our well-being consists in the (1) whole-hearted and (2) successful pursuit of (3) valuable (4) activities."³⁴⁷ In turn, he argues that the existence of a "social form" or "social practice" supporting some activity is, generally, a necessary condition for its value. "[V]aluable activities and pursuits depend on social practices for their availability and, to a degree, even for their existence"³⁴⁸ Raz gives the example of bird-watching:

Bird watching seems to be what any sighted person in the vicinity of birds can do. And so he can, except that that would not make him into a bird watcher. He can be that only in a society where this, or at least some other animal tracking activities, are recognized as leisure activities Much of the interest that people have in goals of these kinds is available to them because of the existence of suitable social forms.³⁴⁹

In short, "[a] person's well-being depends to a large extent on success in socially-defined and determined pursuits and activities."³⁵⁰

The sentence meaning of legal decisions surely is one standard mechanism by which law can influence social practice, specifically the kind of social practice that Raz calls a "social form." In turn, Raz tells us that objective values—objectively valuable activities—cannot be re-

NOZICK, ANARCHY, STATE, AND UTOPIA 42-45 (1974).

³⁴⁷ JOSEPH RAZ, *Duties of Well-Being*, in *ETHICS IN THE PUBLIC DOMAIN* 3, 3 (1994).

³⁴⁸ *Id.* at 18 n.17. In some respects, the views about well-being set forth by Raz in this article are different from the earlier views presented in RAZ, *supra* note 109, at 288-320. The two works are consistent, however, in emphasizing the importance of social forms or practices. I therefore cite the two interchangeably here. See also Joseph Raz, *Mixing Values*, 65 *PROC. OF THE ARISTOTELIAN SOC'Y* 83 (1991) (similarly emphasizing social forms and practices). See generally Roger Crisp, *Raz on Well-Being*, 17 *OXFORD J. LEGAL STUD.* 500 (1997) (presenting a critical analysis of Raz's views, with a particular focus on *The Morality of Freedom*).

³⁴⁹ RAZ, *supra* note 109, at 311.

³⁵⁰ *Id.* at 309.

alized without supporting social forms. Crudely, we could say the following: The degree to which a particular value V_i is realized in a given world-state depends both on whether people successfully perform the relevant activity and on whether the matching social form exists. Whether O^* is better than O with respect to the value of bird-watching is a partial function of whether the social form, bird-watching, exists in O or O^* . To put the point a bit more formally, each objective value has a *cultural* subvariable, such that two world-states identical with respect to social forms fare identically with respect to the subvariable. Given the intimate link between the sentence meanings of A and A^* , and the social forms that exist in O and O^* , is not this Razian objective-list version of utilitarianism very close to a genuine expressive theory?

Close, perhaps, but not close enough. *The linguistic meaning of governmental expression is not the same as its cultural impact.* I emphasized this point repeatedly in Part II while discussing specific expressive theories.³⁵¹ I am now making the same point more abstractly and systematically here. Imagine that governmental action A^* has the sentence meaning "Bird-watching is a valuable activity" and that action A is meaningless. A^* could lead to an outcome in which the social form, bird-watching, is reinforced; it could have no impact; or it could, perversely, undermine the form if, for example, bird-watching is mainly practiced by hippies and "counterculture" types who despise what the government says.³⁵²

The burgeoning literature on law and social norms is of some relevance here.³⁵³ Social norms are not quite the same thing as what Raz calls "social forms." Social norms are a social practice by which activities (existing apart from the practice) are *regulated*. As one leading norms scholar states, "[B]y norms this literature refers to informal social regularities that individuals feel obligated to follow because of

³⁵¹ See *supra* text accompanying notes 199-202, 235-42, 286-90.

³⁵² Technically, given my definition of an expressive factor (as a factor such that meaningless actions necessarily fare the same with respect to it), the example just provided does not prove that a Razian theory is nonexpressive. It is possible that: (1) the status of meaningful A^* versus meaningless A with respect to factor F depends upon further facts beyond the linguistic meaning of the two actions, even though (2) meaningless A and meaningless A' necessarily fare the same with respect to F . See *supra* text accompanying notes 139-40 (discussing different ways in which "expressive factor" could be conceptualized). I have nonetheless used this example, here, for the sake of expository variety; an example of meaningless actions faring differently with respect to the cultural subvariable of Razian objective-list utilitarianism could readily be constructed.

³⁵³ See *supra* note 44 (citing the literature on law and norms).

an internalized sense of duty, because of a fear of external non-legal sanctions, or both."³⁵⁴ Fencing your enclosure, for example, is fencing your enclosure whether or not there is a norm against it. The conditions for a particular behavior to count as "fencing your enclosure" do not include the existence of a no-fencing or pro-fencing norm. By contrast, social forms are a social practice by which activities are wholly or partly *constituted*.³⁵⁵ Bird-watching without the form is not the same as bird-watching with it; that is just Raz's point. Nonetheless, both norms and forms are different kinds of social practice; legal scholars have discussed, at some length, the way in which law affects norms; and this should be at least suggestive for the less examined question: How does law affect social forms?

Notably, the law-and-norms scholarship generally confirms the claim I am making here—that the meaning of a governmental utterance has only a contingent connection to the shape of social practice. A full discussion of the scholarship would take us far afield; I will simply mention the scholars who have developed some kind of general theory of the connection between law and norms. Robert Ellickson, whose seminal book, *Order Without Law*, sparked the entire field, basically views norms as efficient equilibria in iterated prisoners' dilemma games.³⁵⁶ In a one-shot prisoners' dilemma, the equilibrium outcome is socially inefficient. By contrast, in an iterated game, the equilibrium outcome can be efficient, because the threat of retaliation from other players in subsequent rounds can deter individual players from choosing actions leading to an inefficient outcome. Ellickson argues that members of "close-knit group[s] develop and maintain norms whose content serves to maximize the aggregate welfare that members obtain in their workaday affairs with one another"³⁵⁷—norms that are efficient within the group. Law changes these norms by changing the game-theoretic structure upon which the existence of an efficient equilibrium depends:

[T]he number of players involved in an inning of a game, the number of innings in which current players later expect to encounter each other,

³⁵⁴ McAdams, *Origin, Development*, *supra* note 44, at 340.

³⁵⁵ The distinction I am drawing here is, of course, just John Searle's well-known distinction between "regulative" and "constitutive" rules. *See* SEARLE, *supra* note 82, at 33 ("[R]egulative rules regulate antecedently or independently existing forms of behavior, . . . [while] constitutive rules . . . create or define new forms of behavior . . ."); *see also* JOHN SEARLE, *THE CONSTRUCTION OF SOCIAL REALITY* 27-29 (1995) (reiterating this distinction).

³⁵⁶ *See* ELICKSON, *supra* note 44, at 156-83.

³⁵⁷ *Id.* at 167 (emphasis omitted).

the time span within which the players expect those innings to occur, the quality of the players' information, and the distribution of power among the players.³⁵⁸

None of these factors, with the possible exception of the players' "information," has any particularly strong causal connection (let alone a conceptual one) with the meaning of governmental utterances. As Ellickson explains:

Basic rules of land tenure . . . can significantly influence both the number of parties involved in land disputes and the frequency of those parties' encounters. . . . [T]he subdivision of a commons into private parcels abets cooperation by reducing the number of people concerned with localized externalities. . . .

Foundational laws can also lengthen a person's perceptions of the time span For instance, neighbors who own usufructuary interests . . . are likely to have less permanent relationships than neighbors who have life estates³⁵⁹

Even the informational factor, which concerns the information that people end up possessing, should be distinguished from the information that government provides to them (that is, the descriptive content of governmental utterances). For example, Ellickson cautions lawmakers against "imposing new regulatory burdens on the collection and dissemination of truthful, publicly available information about past behavior."³⁶⁰

Robert Cooter has a somewhat different picture of norms—unlike Ellickson, he argues that the existence of a norm involves the punishment of nonconformers by "norm-enforcers" who have "internalized" the norm³⁶¹—but he ends up reaching broadly similar conclusions about the mechanisms by which law changes norms. Cooter points specifically to the role of law in changing the information that group members possess and in changing the close-knit nature of

³⁵⁸ *Id.* at 284. See generally *id.* at 280-86 (summarizing interaction between law and norms).

³⁵⁹ *Id.* at 284-85.

³⁶⁰ *Id.* at 286.

³⁶¹ See Cooter, *supra* note 44, at 953-57. By "internalized," Cooter means, roughly, that the enforcers prefer to conform independent of the threat of external sanction. See *id.* at 956-57. The article cited here is an elaboration of earlier work on norms by Cooter. See *id.* at 953 n.38 (citing Robert D. Cooter, *Decentralized Law for a Complex Economy: The Structural Approach to Adjudicating the New Law Merchant*, 144 U. PA. L. REV. 1643 (1996) and Robert D. Cooter, *The Theory of Market Modernization of Law*, 16 INT'L REV. L. & ECON. 141 (1996)).

groups.³⁶²

Information looms particularly large in the account offered by Richard McAdams, who claims that norms involve a preexisting moral consensus among some portion of the population, which becomes widely known and is then informally enforced by the withdrawal of public esteem from norm-violators.³⁶³ Yet McAdams, like Ellickson³⁶⁴ and Cooter,³⁶⁵ recognizes that government can facilitate or undermine the spread of information required for a consensus to be known, not just by providing the information itself, but also by changing the legal regime governing information acquisition. For example, McAdams argues in favor of a legal regime protecting private, sexual information on the grounds that this kind of regime weakens unwelcome norms governing sexual practices, such as the norm against homosexuality.³⁶⁶

Eric Posner advances quite a different model of norms than do Ellickson, Cooter, or McAdams.³⁶⁷ Unlike Cooter and McAdams, Posner denies that the existence of a norm involves its "internalization" by the group in which the norm exists or some subset thereof, or a consensus among the group or a subset thereof that the disfavored behavior is morally wrong.³⁶⁸ Unlike Ellickson, Posner conceptualizes norms as equilibria in so-called "signaling" games rather than in iterated prisoners' dilemmas.³⁶⁹ (In these "signaling" games, "good types," whose preferences make them well-suited to cooperate with other parties, take actions to signal their preference structure, while "bad types," whose preferences predispose them to cheat those with whom they interact, attempt to mimic the "good types.") Nonetheless, Posner's

³⁶² See Cooter, *supra* note 44, at 972-78.

³⁶³ See McAdams, *Origin, Development*, *supra* note 44, at 355-75.

³⁶⁴ See *supra* text accompanying note 360.

³⁶⁵ See Cooter, *supra* note 44, at 974 ("Informal punishment of transactional wrongs fails when broadcasting [information about the wrong] costs too much relative to the value of a stern reputation. One remedy is to lower the cost of stigmatizing wrongdoers . . ."). This formulation suggests that a (meaningless) legal change that allows private parties to "stigmatize" wrongdoers more freely would be one type of remedy Cooter is envisioning here. See also *id.* at 974-75 & nn.88-89 (arguing that the threat of an employee suit for wrongful discharge undermines an employer's ability to broadcast information about the employee's unsatisfactory performance to prospective employers).

³⁶⁶ See McAdams, *Origin, Development*, *supra* note 44, at 427-32.

³⁶⁷ See Eric A. Posner, *Symbols, Signals and Social Norms in Politics and the Law*, 27 J. LEGAL STUD. 765 (1998).

³⁶⁸ See *id.* at 797 & n.53.

³⁶⁹ See *id.* at 767-72.

model strongly supports the view of the connection between norms and governmental expression for which I am arguing here: governmental actions can create, reinforce, or dissolve norms without those actions themselves possessing linguistic meaning. Posner notes that law can change the equilibria in signaling games by “modify[ing] the cost of sending a signal,” by “modify[ing] the payoffs from cooperation,” and by “modify[ing the] beliefs [of those attempting to distinguish between “good” and “bad” types] about the proportion of types in the population;”³⁷⁰ none of these mechanisms entails that law itself is expressive.

Finally, Larry Lessig, in his work on “social meaning,” stresses that the state has “behavioral” as well as “semiotic” techniques for changing the social meaning of actions. Lessig writes:

We can isolate two such behavioral techniques The first is a regulation designed to *inhibit* a certain behavior that would otherwise aid in the construction or reinforcement of a disfavored social meaning. Segregation, for example, is both an instance of racial harm and a behavior that reinforces the social meaning of inequality. Prohibiting segregation is a way of undermining practices that reinforce social meanings of stigma and inequality.

Another example makes the point more directly. Under the Fair Housing Act, it is illegal for a real estate broker to indicate, whether asked or not, what the racial makeup of a community is when a buyer is purchasing residential property. . . . [Such] restrictions attempt to reduce the number of economic decisions made on the basis of race. . . .

. . . .

. . . The second technique is to *induce* actions that tend either to undermine or to construct a particular social meaning. . . . Political ritual is the easiest case, and *Barnette* [in which the government was prohibited from requiring schoolchildren to stand and salute the flag] serves as a helpful guide through this example.³⁷¹

Lessig’s work may be most directly relevant here, because Lessig insists that his focus is on “social meaning,” not social norms.³⁷² Perhaps this

³⁷⁰ *Id.* at 778; see also *id.* at 789 (summarizing the effect of law on signaling equilibria).

³⁷¹ Lessig, *supra* note 22, at 1013-14. The “semiotic” techniques are generally described *id.* at 1009-12, while the “behavioral” techniques are generally described *id.* at 1012-14.

³⁷² More precisely, Professor Lessig has clarified in subsequent works that the concept of “social meaning” at issue in *The Regulation of Social Meaning*, *supra* note 22, is distinct from the concept of “social norms” as that concept is typically understood by

distinction is better cast as the distinction between social *forms* and social norms. I am skeptical that the kinds of activities Lessig generally focuses upon, such as sexual practices or smoking,³⁷³ possess linguistic meaning;³⁷⁴ more plausibly, the activities are *formalized*, in that such activities are partly constituted by their cultural cast. Smoking is not the same activity as putting wrapped tobacco in your mouth; practicing safe sex is not the same as using prophylactics in a society that lacks sexual rituals.

If this assertion is correct, and Lessig is indeed best construed as analyzing the effect of law on social forms, then Lessig's analysis directly illuminates the question at stake here: Is Razian (social form-based) objective-list utilitarianism a genuine expressive theory of law? The answer, that analysis suggests, is no. Just as the work of Professors Ellickson, Cooter, Posner, and McAdams supports the view that the linguistic meaning of government's utterances has at best a contingent connection to the shape of social *norms*, so the work of Professor Lessig supports the view that the linguistic meaning of government's utterances has at best a contingent connection to the shape of social *forms*.

To sum up: No version of utilitarianism is genuinely expressive. The utilitarian compares two governmental actions, A^* and A , by reference to the total well-being contained in the world-states O^* and O that those actions would produce. More specifically, *preference-utilitarianism* evaluates O^* and O in light of the preferences $P_1 \dots P_n$ of individual persons; *hedonic utilitarianism* evaluates the two states in light of fundamental feeling-types $F_1 \dots F_n$; and *objective-list utilitarianism* evaluates the two states in light of objective values $V_1 \dots V_n$. But neither the P_i , nor the F_i , nor the V_i are genuine expressive factors, because actions A^* and A can fare differently with respect to a P_i , or an F_i , or a V_i , even though neither action is meaningful—even though neither possesses linguistic meaning. This is true notwithstanding the fact that certain variants of utilitarianism may be genuinely *social* or *cultural* in the sense that the comparative worth of O^* and O may be a

other legal scholars, in particular, law-and-economics scholars. See Lawrence Lessig, *The New Chicago School*, 27 J. LEGAL STUD. 661, 680-85 (1998); Lawrence Lessig, *Social Meaning and Social Norms*, 144 U. PA. L. REV. 2181, 2182-86 (1996).

³⁷³ See Lessig, *supra* note 22, at 1019-34.

³⁷⁴ See *id.* at 951 ("Any society or social context has what I call here social meanings—the semiotic content attached to various actions, or inactions, or statuses, within a particular context."). This assertion suggests the linguistic-meaning reading of "social meaning." Lessig, however, quickly goes on to say that his concern is "to find a way to speak of the frameworks of understanding within which individuals live," *id.* at 952, which suggests the social-form reading.

partial function of social or cultural factors, factors that make reference to the social practices (norms, forms) that exist in O^* and O . Preference-based and hedonic utilitarianism are neither genuinely social nor genuinely expressive. Objective-list utilitarianism may be genuinely social, but—given the contingent connection between the linguistic meaning of A^* and A , and the kind of social practices that exist in O^* and O —it is not genuinely expressive either.³⁷⁵

* *

Expressivism will not emerge within utilitarianism, even within its cultural or social variants. Our moral theory must be further enriched. How do we do that?

Let us now consider *consequentialist* theories that are not utilitarian. A moral theory is consequentialist if it has the same general structure as utilitarianism: if it evaluates actions solely in light of the world-states, or outcomes, that the actions produce.³⁷⁶ Utilitarianism is the specific variant of consequentialism that uses the criterion of *overall well-being* to rank world-states. Other, nonutilitarian variants of consequentialism employ other criteria: either they omit the criterion of overall well-being and replace it with some other(s), or they retain that criterion but supplement it with other(s). Does the move within consequentialism from utilitarian to nonutilitarian variants trigger the emergence of genuine expressivism?

³⁷⁵ What about incorporating into utilitarianism an *expressive* theory of individual well-being modeled roughly on Elizabeth Anderson's expressive theory of individual rationality? Anderson herself is hardly a utilitarian, but might it not be possible to employ her theory for utilitarian purposes and to stipulate that government should maximize overall well-being, with each individual's well-being defined as some function of what that person says? Yet even this sort of theory would not be a genuinely expressive theory of law or governmental action. Imagine, for example, that individual lives are better, *ceteris paribus*, to the extent that individuals express love for close friends and family. Factor E_i measures the extent to which that expression of love occurs; the world is better with respect to E_i to the extent that appropriate love is appropriately expressed, and worse, to the extent that it is not expressed (or that inappropriate love is expressed). Yet E_i looks to the totality of what loving and nonloving individuals say about love, not merely to what government says about love or anything else. Thus, depending on the facts, government's saying, "The only thing 'free love' is free of is love" or, "Just say no" might augment, decrease, or have no effect upon E_i . Conversely, a theory that looked essentially to what the government said would not be a utilitarian theory (except if it were based on an utterly implausible view of individual well-being as depending solely on what the government said).

³⁷⁶ See generally sources cited *supra* note 123 (defining and debating consequentialism). A particularly precise definition of consequentialism is provided by McNaughton & Rawling, *supra* note 142.

One plausible source of variation, within consequentialism, concerns the proper distribution of well-being. Utilitarianism employs an aggregative approach: O^* is better than O just in case total well-being in O^* is greater than total well-being in O . Other variants of consequentialism employ nonaggregative approaches. For example, *egalitarian* consequentialism says that O^* is better than O just in case well-being in O^* is spread more equally than well-being in O .³⁷⁷ *Maximin* consequentialism says that O^* is better than O just in case the worst-off person in O^* is better off than the worst-off person in O .³⁷⁸ One hybrid sort of consequentialism says that the comparison of O^* and O is a function both of the total well-being in the two worlds and of the degree to which well-being is spread equally in the two worlds. This sort of variation along the dimension of distribution, however, does not change the nonexpressive cast of consequentialism. Well-being is still what is being distributed, and how each person fares with respect to well-being remains a function of her preferences, her feeling-states, or the objective values she realizes, not the linguistic content of government's actions. For A^* to fare better than A with respect to the total amount of well-being, the equality of well-being, the well-being of the worst-off, or any combination of these factors, it is not necessary that A^* or A be subsumed under a linguistic convention matching actions to sentence meanings.

A second plausible dimension of variation, within consequentialism, concerns the item that is being distributed. What characterizes

³⁷⁷ More precisely, the theory just described is egalitarian welfarist consequentialism. Another variant of egalitarian consequentialism is egalitarian resourcist consequentialism, which is concerned with equalizing primary goods rather than welfare itself. Similarly, the term "maximin consequentialism" that I will use in a moment is technically "maximin welfarist consequentialism." See *infra* text accompanying notes 379-82 (distinguishing between consequentialist views that focus on welfare and consequentialist views that focus on primary goods). On the possibility of egalitarian consequentialism, see KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 48-54. The best discussion of the various egalitarian criteria for ranking outcomes is LARRY TEMKIN, *INEQUALITY* (1993).

³⁷⁸ As Samuel Scheffler explains:

I believe that there are . . . a number of consequentialist conceptions capable of accommodating the objection based on distributive justice. . . . [One such] view identifies the best state of affairs from among a set [as follows]. . . . Given two states of affairs . . . the better state of affairs is the one that maximizes the position of the worst-off group. If the two states of affairs are identical in this respect, the better state of affairs is the one that minimizes the number of people in the worst-off group (by relocating them upward). If the two states of affairs are identical in both of these respects, then the better state of affairs is the one that maximizes the position of the next worst-off group, and so on.

SCHEFFLER, *supra* note 123, at 26-27.

utilitarianism as a distinct moral view is both its aggregative approach and the fact that it aggregates *well-being* rather than something else. John Rawls has famously argued that a moral theory should focus on the distribution of *primary goods*, rather than well-being.³⁷⁹ The nonutilitarian cast of *A Theory of Justice* is bidimensional: Rawls moves (1) from an aggregative to a maximin approach to distribution, and (2) from well-being to primary goods as the basis for determining the position of the least well-off.

The difference principle [i.e., maximin] . . . introduc[es] a simplification for the basis of interpersonal comparisons. These comparisons are made in terms of expectations of primary social goods . . . [which] are things which it is supposed a rational man wants whatever else he wants. . . . With more of these goods men can generally be assured of greater success in carrying out their intentions and in advancing their ends, whatever these ends may be. The primary social goods, to give them in broad categories, are rights and liberties, opportunities and powers, income and wealth. ([Another] very important primary good is a sense of one's own worth . . .)³⁸⁰

Rawlsian primary goods bear particular mention here, given the recurrent role that the primary good of self-respect ("a sense of one's own worth") has played within expressive theories. In Part II, I described the role of *self-respect* within Jean Hampton's expressive theory of punishment, within expressive theories of the Equal Protection Clause, and within the view that the Establishment Clause prohibits the "endorsement" of religion.³⁸¹ Self-respect is linked to social forms and norms—it is plausibly a necessary condition for full self-respect that I lack the social status of an inferior; that I am a first class rather than an *n*th-class citizen—but these social practices, constitutive of self-respect, are conceptually distinct from the linguistic meaning of government's utterances. I made this point, in a general way, when

³⁷⁹ See JOHN RAWLS, *A THEORY OF JUSTICE* 90-95 (1971).

³⁸⁰ *Id.* at 92; see also KOPPELMAN, *supra* note 32, at 205 (discussing Rawlsian self-respect, with specific reference to stigma). Other prominent scholars have pursued Rawls's lead and argued that the "currency" of justice is primary goods, or something close to that, rather than well-being. See G.A. Cohen, *Equality of What? On Welfare, Goods, and Capabilities*, in *THE QUALITY OF LIFE* 9 (Martha Nussbaum & Amartya Sen eds., 1993); Ronald Dworkin, *What Is Equality? Part 2: Equality of Resources*, 10 *PHIL. & PUB. AFF.* 283 (1981); Nagel, *supra* note 154, at 142-72; Thomas Scanlon, *The Moral Basis of Interpersonal Comparisons*, in *INTERPERSONAL COMPARISONS OF WELL-BEING*, *supra* note 339, at 17-44; Amartya Sen, *Capability and Well-Being*, in *THE QUALITY OF LIFE*, *supra*, at 30.

³⁸¹ See *supra* Parts II.A-B.

discussing Razian objective-list utilitarianism.³⁸² The same point holds true in explaining why a nonutilitarian variant of consequentialism that focuses on primary goods rather than welfare is not genuinely expressive. Government does not necessarily create or undermine the self-respect of some persons (the worst off, or anyone else) by valorizing or stigmatizing them, or by saying anything else.

A third plausible dimension of variation, within consequentialism, concerns the function of *desert*.³⁸³ A desert claim takes the following form: a given state or event is deserved by some person in virtue of some feature of hers, such as her prior wrongdoing, her diligent efforts, her free choice, her athletic, musical, or other nonmoral merit, or her moral virtue. To give some standard examples:

1. Jones deserves his success; he's worked hard for it.
-
3. Walters deserves the job; he's the best-qualified applicant.
4. Wilson deserved to be disqualified; he knew the deadline for applications was March 1.
5. Jackson deserves more than minimum wage; his job is important and he does it well.
6. Baker deserves to win; he's played superbly.
-
8. [Applebaum] deserves his twenty-year sentence; he planned the murder.
-
10. Winters deserves some compensation; he's suffered constant pain since the shooting.
11. Lee deserves a reward; he risked his life.
-
13. Gordon deserves some good luck; he's had only bad.³⁸⁴

An expressive account of desert might in theory stipulate either that the desert-basis is sometimes expressive, or that the deserved outcome is sometimes expressive, or both. In all of the above examples, however, the desert-basis is nonexpressive—what matters is that the person worked hard, possessed the best qualifications, ignored the predictable outcome, executed an important job, performed a meritorious action, committed culpable wrongdoing, and so forth, not that she

³⁸² See *supra* text accompanying notes 351-75.

³⁸³ See KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 54-59 (discussing the possibility of injecting considerations of desert into consequentialism).

³⁸⁴ SHER, *supra* note 191, at 7. Sher uses "Anderson," not "Applebaum," in the original, but since I refer to Professor Elizabeth Anderson at various points in the Article, I use "Applebaum" here to avoid confusion.

said something—and in any event the expressive cast of the desert-basis does not make the case for an expressive theory of *law*. Government's role, if any, is in meting out deserved states or events. So let us focus on the kind of expressive, desert-based theory that construes the deserved outcome to be expressive.

On this kind of theory, the occurrence of some desert-basis (such as hard work, good qualifications, voluntary risk-taking, practical accomplishment, meritorious action, or culpable wrongdoing) improves the consequential evaluation of certain *statements*. Those statements are ascribed greater value than if the desert-basis had not occurred. For example, the statement "We praise Baker, who was the best pianist in this competition" might be accorded intrinsic value, conditional upon Baker's actually being the best pianist in the group. If he is, a laudatory statement is "deserved" and counts as a good thing quite apart from any further beneficial consequences the statement may have. Similarly, the statement "Applebaum is condemned" might be accorded intrinsic value, conditional upon Applebaum's being a wrongdoer—this is, in effect, what Robert Nozick's expressive and desert-based theory of punishment does.³⁸⁵ Or, a statement such as "Winters deserves to be compensated by Fred" might become intrinsically valuable in virtue of the fact that Fred injured Winters. A statement such as "Jones is a success!" might become intrinsically valuable in virtue of Jones's hard work.

One problem with this kind of view is that it has limited scope; it is intuitively plausible for the cases of meritorious action and culpable wrongdoing, but not for the further cases delineated in the above list of examples.³⁸⁶ What the injured Winters, intuitively, deserves is not a statement, but whatever it takes to redress the injury that Fred caused to him. Likewise, what Jones deserves is success; what Walters deserves is the job; what Wilson deserves is disqualification. Another problem is that the expressive account of desert is ad hoc, even within the limited domain of meritorious and wrongful action. Applebaum does, intuitively, deserve condemnation and not merely hard treatment—to return to the discussion in Part II, he deserves *Punishment_{E+H}*, not merely *Punishment_H*—but, as we have seen, existing attempts to justify

³⁸⁵ See *supra* text accompanying notes 187-88 (discussing Nozick's theory).

³⁸⁶ Arguably, the account has intuitive plausibility for the case of meritorious *character* as well as meritorious action. See SHER, *supra* note 191, at 132-49 (analyzing desert claims that arise from virtue). But if so, the expressivist account of the meritorious character case is subject to my criticisms of expressivism regarding the desert-claims of bad and good actors.

this intuition are unpersuasive. Robert Nozick argues that wrongdoers deserve *Punishment_{E+H}* not merely *Punishment_H* because *Punishment_{E+H}* tracks wrongdoing. I suggested in Part II that the argument was fallacious. Poetic justice (the landslide falling on the murderer) does not track wrongdoing, but a reliable system of *Punishment_H* would do so, even granting Nozick his controversial premise that desert entails tracking.³⁸⁷

What about an expressive, desert-based account of meritorious action? Given the rough symmetry between meritorious and culpable action, it is hard to see why that account should succeed where an expressive account of punishment and wrongdoing fails. If wrongdoers do not, in fact, deserve condemnation—and no one has yet shown that they do—then why assume that good-doers deserve praise?³⁸⁸ At most, government and private actors are under generic requirements of veracity and justification that apply, *inter alia*, to cases of culpable wrongdoing and meritorious action.³⁸⁹ If Applebaum deserves *Punishment_H* (hard treatment), and if, further, government has a generic moral reason to furnish a justification for its decisions, then government has a moral reason both to inflict hard treatment upon Applebaum and to utter a statement describing his desert-basis (the fact of his culpable wrongdoing). Similarly, if government has a generic reason not to lie, and if Baker really was the best pianist in the competition, then government has a moral reason not to award the prize for “Best Pianist” to anyone other than Baker (insofar as the conferral of this award upon anyone else involves the false assertion that the recipient was better than Baker). As George Sher explains: “By . . . appealing to veracity, we can justify the desert-claims that arise when persons display merit in conventionally structured contexts.”³⁹⁰ He elaborates:

If awarding a prize or grade in accordance with fixed rules amounts to making an assertion about how well the recipient has performed, then deviating from those rules must amount to making a *false* assertion about this. Hence, whatever is wrong with making such false assertions must also be wrong with departing from the rules. . . . [I]f someone who promulgates, referees, or judges a contest knowingly deviates from the

³⁸⁷ See *supra* text accompanying notes 187-88 (criticizing Nozick’s expressive account of punishment).

³⁸⁸ See SHER, *supra* note 191, at 111 (noting the symmetry between desert-claims of wrongdoers and meritorious actors).

³⁸⁹ See *id.* at 115-18 (considering the extent to which the principle of veracity explains desert-claims of meritorious actors).

³⁹⁰ *Id.* at 117.

rules, and thus makes (or contributes to the making of) a false assertion about a performer's merit, he arguably does violate his obligation of veracity.³⁹¹

Note, however, that the generic "expressive" duties of truth-telling and justification are not limited to the case of a deserving actor—these obligations would come into play even where no desert-basis has occurred—and that the statement required or proscribed by these duties is a *descriptive* statement, not the statement of praise or condemnation purportedly triggered by merit or wrongdoing.³⁹²

* *

We have thus far considered two types of moral enrichment that might be thought to trigger the emergence of an expressive theory:

³⁹¹ *Id.* at 116 (citation omitted).

³⁹² Cass Sunstein suggests that expressive considerations may enter consequentialism via the role of "expressive norms" in ranking states of affairs.

We might be tempted to suppose that people can avoid expressive concerns entirely and that it is possible to assess law solely on the basis of consequences But this is not actually possible. The effects of any legal rule can be described in an infinite number of ways. Any particular characterization or accounting of consequences will rest not on some description of the brute facts; instead it will be mediated by a set of (often tacit) norms determining how to describe or conceive of consequences. It is possible to see a large part of the expressive function of law in the identification of what consequences count

Sunstein, *On the Expressive Function of Law*, *supra* note 20, at 2048. This seems to be a kind of metaethical expressivism: whether outcome *O** is better than *O* depends upon whether, in light of existing norms-for-conceiving-consequences (norms for the practice of moral evaluation), *O** is better. I doubt that Sunstein really means to espouse metaethical expressivism, and in any event there are strong arguments against such a view. If the goodness or badness of consequences has moral weight (*prima facie* or conclusive), then morally good or bad consequences are not simply the consequences that we all approve as good or bad (in light of norms, or otherwise). See *supra* text accompanying notes 61–68 (arguing against noncognitivism and against the noncognitivist construal of Sunstein and other legal expressivists). But even leaving this point aside—even assuming that the moral goodness or badness of *O* and *O** is, at bottom, reducible to their status pursuant to social norms-for-conceiving-consequences—it seems incorrect that *this* status, in turn, has any conceptual connection to the linguistic meaning of governmental utterances. For example, a particular activity might be socially disfavored (both as a matter of general belief, and pursuant to the social norms-for-conceiving-consequences, such that the flourishing of the activity is not conventionally counted as a world-improvement), notwithstanding the efforts of a (progressive or reactionary) government to speak out in favor of that activity. No metaethical expressivist I know of thinks that "X is morally required/prohibited/permitted/good/bad" reduces to "X is legally required/prohibited/permitted/good/bad" or to "government says that X is required/prohibited/permitted/good/bad."

first, the enrichment that occurs *within* utilitarianism when we move from a preference-based view of well-being to other, more sophisticated views; second, the move *from* utilitarianism to a more nuanced kind of consequentialism, which looks not merely to the maximization of well-being but in addition (or alternatively) to considerations of equality, primary goods, and desert. But expressivism does not yet seem to have emerged. Perhaps, then, the best way to construct an expressive theory is by rejecting consequentialism entirely: by moving from a theory that enjoins agents to maximize good consequences (however defined), to a theory that incorporates deontological constraints, such that agents are proscribed from breaching these constraints even where their breach would produce good consequences overall.³⁹³

Deontological constraints do indeed surface within the literature on expressive theories of law. For example, as we have seen, Pildes and Sunstein attempt to defend an expressive theory of regulation by appealing to deontological constraints:

[T]here is a second ground for endorsing the expressive function of law, and this ground is not about social effects . . . [S]omeone might refuse to kill an innocent person at the request of a terrorist, even if the consequence of the refusal is that many more people will be killed. Our responses to this case are not adequately captured in purely consequentialist terms.³⁹⁴

³⁹³ For general discussions of the nature and plausibility of deontological constraints, see KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 70-152; NOZICK, *supra* note 346, at 26-53; SCHEFFLER, *supra* note 123, at 80-114; Richard Brook, *Is Smith Obligated that (She) Not Kill the Innocent or that She (Not Kill the Innocent): Expressions and Rationales for Deontological Constraints*, 35 S.J. PHIL. 451 (1997); F.M. Kamm, *Non-Consequentialism, the Person As an End-in-Itself, and the Significance of Status*, 21 PHIL. & PUB. AFF. 354 (1992); McNaughton & Rawling, *supra* note 142; and Nagel, *supra* note 154, at 156-67.

³⁹⁴ Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 69. Pildes and Sunstein describe this scenario with language suggestive of deontological *options*, rather than constraints. See *id.* ("[I]ndividual behavior is not concerned solely with states of affairs, and . . . if it were, we would have a hard time making sense of important aspects of our lives. Personal integrity, commitment, and the narrative continuity of a life matter enormously as well." (footnote omitted)); see also SCHEFFLER, *supra* note 123, at 41-79 (discussing and defending deontological options); Nagel, *supra* note 154, at 150-56 (same). But the particular example of killing one to save five is generally seen to implicate constraints rather than options, see, e.g., SCHEFFLER, *supra* note 123, at 80-114 (discussing "agent-centered restrictions," i.e., constraints); and in any event, I think it much more plausible that government is subject to deontological constraints than that government is entitled to deontological options, see *supra* text accompanying notes 153-58 (arguing that government lacks its own "projects," of the kind that options protect).

The deontological constraint invoked here—the prohibition on killing one person, even to save more than one person—is indeed a plausible element of morality that cannot be captured within a consequentialist framework and that plausibly applies to governmental actions, not just the actions of private individuals. Government is constrained from killing an innocent person, even if that action will avert a greater number of deaths or killings overall or is consequentially warranted in some other way.³⁹⁵

Note, however, that the existence of a deontological constraint covering governmental actions does not yet produce an expressive theory of government or of lawmaking—it simply produces a nonconsequentialist theory. A deontological constraint simply identifies a particular type of action, the performance of which is prohibited notwithstanding better consequences overall. Unless the constrained action is characterized in linguistic terms, deontology is no more an expressive theory than simple or sophisticated consequentialism. For example, imagine that the prohibition on killing an innocent person is best captured by a constraint on intentional, serious, nondefensive, and undeserved harming.³⁹⁶ Each agent is generally constrained from intentionally producing a significant setback to another person's interests, unless the harm is deserved or the agent acts in self-defense; and the requirement not to (intentionally) kill an innocent person is seen as a specific application of this more general constraint. Such a constraint is nonexpressive because the prohibited actions are characterized in terms of nonlinguistic properties: what matters about an action is whether it causes serious harm (and if so, whether it was performed nonintentionally, defensively, or by virtue of the harmed person's desert), not what the action *means*. The utterance of a statement might constitute a prohibited harming (if, for example, a governmental official utters a statement directing a subordinate to shoot the victim), but so might the performance of a meaningless action (if, for example, the official himself shoots the victim), and, clearly, it is quite possible that two meaningless actions will fare differently with respect to the no-harming constraint.

In short, the existence of deontological constraints is useful for

³⁹⁵ See, e.g., KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 71 ("Intuitively, at least, most of us have little doubt that it is morally forbidden to chop up an innocent person, even if this is the only way to save five other innocent people from death. Some acts are morally off-limits; they are forbidden, even if the results would be good.").

³⁹⁶ Cf. *id.* at 70-105 (discussing how a prohibition on killing one to save five might be specified in light of distinctions between doing and allowing harm, intending and foreseeing harm, defensive and offensive harm, and/or related distinctions).

the expressivist only if she can further show that such constraints (or at least some of them) are “expressive factors” in the sense that I described in Part I—only if she can further show that at least some of the constraints characterize actions by reference to linguistic meaning. Sunstein and Pildes do not attempt to make this further showing; they simply, and inconclusively, point to the existence of deontological constraints.³⁹⁷ Similarly, Dan Kahan leaps from the fact that nonconsequential gradations of wrongdoing exist to the conclusion that an expressive theory of wrongdoing is correct.

[C]ompare the actions of a white supremacist who kills an African-American out of racial hatred and a mother who in anger kills a man who has sexually abused her child. Both acts are wrong, and their consequences are in some sense equivalent—there is one dead person in each case. Nevertheless, the racist’s killing is more worthy of condemnation precisely because his hatred expresses a more reprehensible valuation than does the mother’s anger.³⁹⁸

Assume Kahan is correct that no consequentialist account of the difference between the white supremacist and the avenging mother is available.³⁹⁹ He still needs an argument that the best deontological account of the difference is an expressive account. An expressive constraint could explain the difference—perhaps there is a special constraint against making insulting and degrading statements, which is breached by the supremacist but not by the mother⁴⁰⁰—but alternatively, and more simply, one might say that the supremacist has both killed the victim and damaged the status and self-respect of other blacks. On this alternative account, a nonexpressive action that was equally harmful as the supremacist’s expressive action, in terms of the death and/or status harms the two actions produced, would be an equal deontological wrong.⁴⁰¹ My aim here is not to defend this par-

³⁹⁷ See Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 69.

³⁹⁸ Kahan, *supra* note 9, at 598.

³⁹⁹ It is far from clear that no consequentialist account is persuasive. For example, there might be a contingent, causal connection between the expressive content of the supremacist’s action and the status and self-respect of blacks; self-respect, in turn, could well play a role within a consequentialist moral theory. See, e.g., *supra* text accompanying notes 379-82 (discussing the link between stigma, status, and self-respect in the context of nonutilitarian consequentialism).

⁴⁰⁰ This is, in effect, what Jean Hampton argues. See Hampton, *supra* note 28, at 1671-85 (developing an expressive account of the distinction between wrongdoing and mere harm-doing).

⁴⁰¹ More precisely, an action that was equally harmful in these or other deontologically significant ways and that had the further, nonexpressive characteristics needed to trigger a deontological constraint, e.g., the requisite intentionality on the actor’s part,

ticular, alternative account of Kahan's example, but rather to make the conceptual point that deontology has nonexpressive variants as well as the expressive variants presupposed by Kahan, Pildes, and Sunstein.

Indeed, the standard theory of deontology is, centrally, a nonexpressive theory. Return to the central and paradigmatic case that deontological theorists have focused upon—the case of killing one to save five. Two purported constraints are standardly offered to cover this case. The first is a constraint on *intending* harm (with further qualifications to cover self-defense, deserved harm, and so forth), such that harmful actions or omissions which are merely foreseen but not intended are not covered by the constraint. The second is a constraint on *doing* harm (again with suitable qualifications), such that harmful actions, foreseen or intended, are prohibited, but harmful omissions, even intended ones, are not.⁴⁰²

The basic moral factor of harming, shared by both proposals, is not an expressive factor, nor are the further factors of intentionality and act/omission. Given two meaningless actions, one can be harmful and the other not. Given two meaningless actions, one can be intentional and the other not. Given two meaningless behaviors, one can constitute an action and the other an omission.

Nor is it plausible to think that the standard proposals can be replaced with a constraint that makes expression a necessary condition for a deontological violation. The view that deontological constraints prohibit *only* actions with a given meaning is clearly mistaken. At certain points, Jean Hampton seems to adopt such a view. For example, she states that “[i]t is because behavior can carry meaning with regard to human value that it can be wrongful.”⁴⁰³ If Hampton means to claim, here, that a meaningless action cannot be nonconsequentially wrongful, then she is clearly incorrect. Consider the action of intentionally shooting an innocent victim at close range, where the actor neither implicates a sentence-meaning convention nor, in fact, possesses a speaker's meaning. Surely the occurrence of such an action is logically possible: sentence-meaning conventions might not be triggered; the actor might not possess the further mental states, beyond a

would be deontologically wrong and equally so; an equally harmful action that did not have those further characteristics would not be deontologically wrong; and a differentially harmful action with those further characteristics would be a differentially serious deontological wrong.

⁴⁰² See KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 94-105 (presenting and analyzing these two proposals).

⁴⁰³ Hampton, *supra* note 28, at 1670.

bare intention to kill, necessary to constitute a speaker's meaning; the action might directly cause the innocent victim's death. Surely, too, if any deontological constraints exist, then one covering this action does.

This leaves the options of (1) retaining the standard no-harming constraint but building in an expressive factor, such that a harmful action is worse if it says the wrong thing or (2) creating a second and separate expressive constraint, such that a harmless action (one permissible under the main no-harming constraint) can violate this second one if the action says the wrong thing. Indeed, there is a fine example of these options in the philosophical literature: the irreducible prohibition against lying, understood either in the form of (1), as the proposition that a lie makes intentional or active harming worse than such harming would be otherwise; or in the form of (2), as an independent constraint against lying even if the lie is harmless.⁴⁰⁴ Either way, an expressive feature of actions—the fact that the action is a lie, that it has a descriptive meaning that turns out to be false—is proposed to have genuine, deontological significance.⁴⁰⁵

Is this proposal persuasive? Is there truly an irreducible prohibition on lying, or rather just the basic constraint on intentional or active harming, applicable both to lying and to nonmeretricious harming? As Shelly Kagan explains:

We certainly know that in the *typical* case telling a lie is harmful. . . . Knowing the truth is an extraordinarily valuable means of achieving one's goals. When you tell a lie, you virtually always rob someone of a crucial ingredient—information—that they need to accomplish whatever it is that they wanted to accomplish Sooner or later, the person is likely to act on the misinformation you've given them, or fail to act when they should have acted, and the result will typically be frustration and disappointment So if we already have a constraint against doing

⁴⁰⁴ For a discussion focused on the possibility of a deontological constraint against lying, see CHARLES FRIED, *RIGHT AND WRONG* 54-78 (1978), and KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 106-16. For more general discussions of the moral status of lying, see F.G. BAILEY, *THE PREVALENCE OF DECEIT* (1991); SISSELA BOK, *LYING: MORAL CHOICE IN PUBLIC AND PRIVATE LIFE* (1978); CHRISTINE M. KORSGAARD, *CREATING THE KINGDOM OF ENDS* 335-62 (1996); and Alan Strudler, *Incommensurable Goods, Rightful Lies, and the Wrongness of Fraud*, 146 U. PA. L. REV. 1529 (1998).

⁴⁰⁵ Actually, we can debate whether the anti-lying constraint is genuinely expressive. If, for example, it prohibits all actions designed to cause false beliefs in some target, then the constraint is nonexpressive, because agents can intentionally cause false beliefs through nonlinguistic means. But I will ignore this point and assume that the anti-lying constraint, if it exists, is indeed expressive, by constraining only *linguistic* utterances that are untrue or cause false beliefs.

harm, we can easily derive from it a [reducible] constraint against telling lies.⁴⁰⁶

It is well beyond the scope of this Article to take a stance on the deontological status of lying. Note, however, that even if an irreducible prohibition on lying does exist, that prohibition supports a very different kind of expressive theory than those advanced to date in the legal literature. First, the theory is generic: it applies to all official actions, not merely actions of a particular kind of legal official or actions subsumed within a particular area of law. (This is by contrast with the expressive theories of punishment, the Equal Protection Clause, and the Establishment Clause, all of which posit constraints or requirements considerably narrower than the generic no-lying constraint.) Second, it is a prohibitory theory rather than a mandatory one. The no-lying constraint does not oblige official actors to perform certain utterances, but merely, and less demandingly, to refrain from certain ones, namely lies.⁴⁰⁷ (This is by contrast with the expressive theory of punishment, which demands an affirmative act of condemnation in response to wrongdoing, and the expressive theory of regulation, which demands an affirmative statement describing the values that bear upon a regulatory choice.⁴⁰⁸) Third, the no-lying constraint is focused on descriptive meaning rather than on directive, declarative, expressive, or commissive meaning. (This is by contrast with the expressive theories of punishment and of the Establishment Clause, which focus, respectively, on the expressive acts of condemnation and endorsement, and, perhaps, with the expressive theory of the Equal Protection Clause, if "stigma" is understood as a declarative rather than descriptive type of meaning.⁴⁰⁹)

I find it highly implausible that the sheer action, by a legal official, of uttering a *prescriptive* statement—that is, the sheer action of creating legal rights, duties, liabilities, immunities, or other legal positions—routinely possesses a further descriptive content beyond whatever descriptive content inheres in its prescriptive meaning. I also find it highly implausible that the no-lying constraint, such as it may be, is

⁴⁰⁶ KAGAN, *NORMATIVE ETHICS*, *supra* note 123, at 108.

⁴⁰⁷ See SHER, *supra* note 191, at 112.

[O]n most interpretations, th[e] principle [of veracity] is negative rather than positive. It requires that we not say what we know to be false, but it does not require that we say everything we know to be true. Unless others have special rights to our information, usually we may simply remain silent.

Id.

⁴⁰⁸ See *supra* text accompanying notes 162, 300-01, 332-33.

⁴⁰⁹ See *supra* text accompanying notes 162, 223, 249-55.

implicated by whatever descriptive meaning does inhere in prescriptive meaning. (Every legal utterance may include the descriptive rider, "It is morally required that ____," and that rider may be routinely false,⁴¹⁰ but the upshot cannot be that our legal system must grind to a halt—that officials must cease creating legal rights, duties, and so on.) This suggests that the no-lying constraint, albeit generic, is not really a constraint on lawmaking at all. Rather, it is a constraint on the *further* utterances of all governmental officials—on their utterances that lack prescriptive meaning altogether and that merely describe. In effect, it is a constraint on the contents of the Congressional Record and the House and Senate Reports (not the U.S. Code), on the contents of the Federal Register (not the Code of Federal Regulations), and on the contents of the judicial reporters (not the legal orders that judges issue to litigants). A moral theory incorporating the no-lying constraint is, indeed, a genuinely expressive theory, but it is not yet a genuinely expressive theory of *law*.⁴¹¹

To be sure, it could be the case that morality incorporates multiple expressive constraints: both the no-lying constraint plus additional constraints that significantly limit prescriptively meaningful utterances. This is, however, an esoteric view of deontology—the only expressive constraint that deontologists standardly accept is the no-lying constraint⁴¹²—and there is reason to doubt that the view can be successfully defended.⁴¹³ Deontological constraints are rooted in the separateness of persons—in the fact that the would-be victim, protected by the constraint, has her own life to lead and should not be sacrificed even for the sake of the greater good.

The victim feels outrage when he is deliberately harmed even for the greater good of others, not simply because of the quantity of the harm but because of the assault on his value of having my actions guided by his evil. . . .

The five people I could save by killing him can't say the same, if I re-

⁴¹⁰ See *supra* text accompanying notes 111-12 (discussing possible conceptions of prescriptive meaning to accommodate the Razian point that law claims *de facto* authority).

⁴¹¹ Might a similar objection hold good against all of the other variants of expressivism that I have analyzed in this Article? I have not generally pursued the issue with respect to these variants because I have generally advanced the logically prior, and considerably more robust, claim that the expressive factors posited by these variants do not exist.

⁴¹² See Nagel, *supra* note 154, at 157 (listing constraints).

⁴¹³ For one attempt to defend an expressive constraint in addition to the no-lying constraint, see Hampton, *supra* note 28, at 1671-85.

frain. They can appeal only to my objective acknowledgment of the impersonal value of their lives. That is not trivial, of course, but it still seems less pressing than the protest available to my victim . . . as the possessor of the life I am aiming to destroy.⁴¹⁴

This motivates the main deontological constraint against harming, and perhaps the constraint against lying—insofar as lying, quite apart from harming, is understood as a special breach of the victim's autonomy. Additional factors (for example, intentionality or the action/omission distinction) are brought into play so as to limit the scope of the main constraint and give the actor adequate space to live her own life, or, conversely, to pick out those harmful actions where her agency is distinctly involved. Beyond all this, does the separateness of persons license an additional expressive constraint? Why would it? The linguistic meaning of an action is linked only contingently to its harmfulness (even in the oft-used case of stigmatic harm); linguistic meaning may be essentially linked to an infringement of autonomy,⁴¹⁵ but the no-lying prohibition would seem to exhaust that consideration; and it is quite obscure how the need to protect the actor's own projects, or to highlight her agency, would give rise to expressive factors. The considerations of harm, autonomy, and agency that underlie deontological constraints are not, I think, likely to prove a fertile ground for expressivism.

* *

I now turn, briefly, to a final source of enrichment of moral theory: the introduction of *political* considerations into moral theory. By this, I mean those considerations that pertain exclusively to government or to other political associations.⁴¹⁶ The dimension of enrich-

⁴¹⁴ Nagel, *supra* note 154, at 167. Notably, Nagel is here defending his particular version of the no-harming constraint—incorporating the distinction between intentional and merely foreseen harms, rather than between acts and omissions—but presumably any defender of a no-harming constraint will offer a similar defense (one that appeals in some way to the separateness of persons).

⁴¹⁵ See KORSGAARD, *supra* note 404, at 358 ("If we value the Kantian ideal of free and non-manipulative relations among rational beings, . . . we must learn to be truthful and straightforward with one another . . .").

⁴¹⁶ To be sure, as I have emphasized throughout this Article, any expressive theory of law will posit factors that apply to the actions of legal (or other governmental) officials. The fact, if it is one, that individuals rationally or morally act to satisfy certain expressive (linguistic) requirements does not yet show that government does. See *supra* Part I.E (distinguishing expressive theories of individual action from expressive theories of governmental action). One response is to argue that the moral requirements

ment, here, is orthogonal to the consequentialism-deontology dimension. Political considerations may mean distinctive consequentialist considerations (such that the outcomes that government should promote are distinct from the outcomes that ordinary individuals should promote), distinctive deontological constraints (additional to those governing private actors), or both.

There is a burgeoning scholarly literature, exemplified by John Rawls's book, *Political Liberalism*, on the existence and shape of political considerations.⁴¹⁷ This literature focuses specifically on the special moral obligations attendant upon government in virtue of the plurality of moral and evaluative views held by citizens. As Rawls puts it:

How is it possible that there may exist over time a stable and just society of free and equal citizens profoundly divided by reasonable though incompatible religious, philosophical, and moral doctrines? . . . How is it possible that deeply opposed though reasonable comprehensive doctrines may live together and all affirm the political . . . regime?⁴¹⁸

The literature on this problem has had relatively little impact on the debate about legal expressivism. For the most part, expressivists have defended their views with reference to moral concepts (such as desert, stigma, or status) that figure in straight moral theory or at least have close analogues there.⁴¹⁹ It is at least worth asking, however, whether Rawlsian or quasi-Rawlsian concepts such as "public reason," "liberal neutrality," or "deliberative democracy"—concepts that truly are unique to governmental action—might prove a fruitful basis for ex-

applicable to individual actors (for example, deontological norms or the injunction to maximize good consequences) also apply to legal officials or other governmental actors. That is, in effect, what the expressive theories considered up to this point do. A different approach is to suggest that there are moral requirements *distinctive* to governmental actions—what I am here calling "political considerations"—and that such requirements are expressive.

⁴¹⁷ JOHN RAWLS, *POLITICAL LIBERALISM* (1993). One part of this literature is comprised of direct critical responses to Rawls. For some examples, see the symposia on *Political Liberalism* in 94 COLUM. L. REV. 1813 (1994); 105 ETHICS 4 (1994); 75 PAC. PHIL. Q. 165 (1994); and 7 RATIO JURIS 267 (1995). Another part is comprised of works of normative theory that, like *Political Liberalism*, adduce "political considerations" as I have defined them here. See, e.g., BRUCE A. ACKERMAN, *SOCIAL JUSTICE IN THE LIBERAL STATE* (1980); GERALD F. GAUS, *JUSTIFICATORY LIBERALISM: AN ESSAY ON EPISTEMOLOGY AND POLITICAL THEORY* (1996); KENT GREENAWALT, *PRIVATE CONSCIENCES AND PUBLIC REASONS* (1995). See generally FRED D'AGOSTINO, *FREE PUBLIC REASON: MAKING IT UP AS WE GO* (1996) (surveying this literature).

⁴¹⁸ RAWLS, *supra* note 417, at xviii.

⁴¹⁹ For one exception, see Pildes, *Why Rights Are Not Trumps*, *supra* note 4, at 735-36 (defending an expressive account of constitutional rights, grounded in part on the proposition that constitutional "rights are the means for enforcing the differentiations of political authority characteristic of liberal societies").

pressivism.

I doubt they do. Imagine, as Rawls posits, that citizens hold different and "reasonable" moral views. Why would this fact of reasonable moral pluralism give rise to expressivism about governmental action, where straight moral reasoning does not? Note, to begin, that expressive factors will not emerge within political theory by virtue of a simple convergence between the differing moral views.⁴²⁰ One person holds reasonable moral view M_1 , while another person holds reasonable moral view M_2 . M_1 is nonexpressive; it includes no expressive factors. Is it unreasonable by virtue of that? Not if the arguments to this point have been cogent. I have argued that morality (apart from the problem of accommodating plural views) is in fact nonexpressive.⁴²¹ *A fortiori*, a nonexpressive moral view is *reasonable*. Therefore, M_1 and M_2 will not converge upon expressivism. This is true even if M_2 itself is expressive—if expressivism is, to some extent, a reasonable view for citizens to hold. The proposition that the factors applicable to governmental action are partly expressive may be part of M_2 , but it is not part of M_1 , and so that proposition is not one upon which the two views will converge.

What about consensus rather than convergence? M_1 might not directly support expressivism, but the holder of M_1 might come to understand that—given the very fact of disagreement with M_2 about the truth of expressivism, or about other matters—the best consensus view between M_1 and M_2 is in some way expressive. Call this consensus view M^* . We now need an argument for the proposition that M^* is expressive notwithstanding the nonexpressive cast of M_1 . This is where the idea of neutrality comes into play.⁴²² Plausibly, M^* includes the requirement that government remain neutral between M_1 and M_2 . In turn, neutrality might be understood in expressive terms. It might be cashed as a constraint on governmental endorsement of M_1 or M_2 .

As we saw in Part II, this is how Donald Beschle construes neutrality in the context of the Establishment Clause. The Clause, he claims, prohibits government from uttering *statements* to the effect that one religion is true, correct, or governmentally approved. The problem with Beschle's construal of Establishment Clause neutrality is that the

⁴²⁰ On the distinction between political theories that posit a *convergence* between differing citizen views and theories that describe citizens as reaching a *consensus*, see D'AGOSTINO, *supra* note 417, at 30-33.

⁴²¹ See *supra* text accompanying notes 337-415.

⁴²² Neutrality looms large in Ackerman's version of political liberalism. See, e.g., ACKERMAN, *supra* note 417, at 10-12, 349-69.

nonexpressive transfer of resources from one church to another, a transfer (let us assume) that is justifiable only in light of the transferee church's doctrine, counts as perfectly neutral and permissible.⁴²³ Generalizing, an expressive conception of the neutrality constraint in M^* will permit government to take any actions (even those exclusively supported by M_1 and wholly at odds with M_2 , or vice versa) that do not have the linguistic meaning of endorsing, favoring, or supporting M_1 or M_2 . That seems much too lax.

In Part II, I sketched out a competing and nonexpressive conception of neutrality: governmental decisions must be justifiable relative to a suitably constrained set of moral propositions, excluding some (or, at the limit, all) of the moral propositions upon which M_1 and M_2 disagree, and in particular excluding the proposition that M_1 or M_2 is true.⁴²⁴ This conception precludes governmental endorsement of M_1 or M_2 , but it also properly precludes nonexpressive actions that are unduly biased towards M_1 or M_2 . For example, if M_1 contains one view of the good life, and M_2 contains another,⁴²⁵ then M^* plausibly prohibits government from coercing the holder of M_2 to follow the life-plan sketched out by M_1 , even if that view is more likely to be correct than the view contained by M_2 .⁴²⁶ This constraint will apply to governmental actions vis-à-vis the holder of M_2 that have no (nonprescriptive) linguistic content, such as the straight, nonexpressive coercion of the holder of M_2 to follow the M_1 life.⁴²⁷

Neutrality is not the only idea that might be used in the

⁴²³ See *supra* notes 273-77 and accompanying text (presenting and criticizing Beschle's view).

⁴²⁴ See *supra* text accompanying note 278.

⁴²⁵ The purported requirement that government remain neutral between different conceptions of the good life has been particularly emphasized by political theorists in recent years. For an overview and critique, see SHER, *supra* note 343.

⁴²⁶ This is so even if, on the M_1 view, the holder of M_2 is made better off notwithstanding the coercive intervention. On some views of welfare, coercive interventions can be welfare-improving. Cf. Joel Feinberg, *Legal Paternalism*, in PATERNALISM 3, 5 (Rolf Sartorius ed., 1983) ("What justifies the absolute prohibition of interference in primarily self-regarding affairs—if there is one—is not that such interference is self-defeating and likely (merely likely) to cause more harm than it prevents, but rather that it would itself be an injustice, a wrong . . ."). The point here is that, whether or not M_1 is the kind of view that counts the coercive intervention as welfare-improving for the M_2 holder, M^* prohibits the intervention.

⁴²⁷ Even if this particular variant of the neutrality constraint is incorrect, the point remains that the right conception—whatever it is—will surely recognize that a governmental action without a biased or nonneutral linguistic meaning can still be biased or nonneutral.

construction of M^* . Alternately, one might appeal to the idea of a hypothetical contract.

The requirement that government enjoy the consent of the governed is deeply rooted in our political culture and is arguably central to liberalism. The most powerful and systematic elaboration of this requirement in recent years has come from liberals who are also contractualists. Contractualist liberals argue that governmental arrangements are justified only if they could or would be accepted by signatories to a hypothetical contract.⁴²⁸

Perhaps the holder of M_1 and the holder of M_2 , under the right conditions, would reach a contract M^* that includes expressive constraints on government notwithstanding the absence of such constraints in M_1 . A full argument to the contrary would be beyond the scope of this Article; it would require a specification of the contracting scenario, something that remains much disputed in the political theory literature.⁴²⁹ Notably, however, the scholars (such as Rawls or Scanlon⁴³⁰) who have undertaken the hypothetical-contract exercise in detail have not ended up with expressive conceptions of M^* . For example, Rawls's M^* is fundamentally concerned with the distribution of primary goods and with the scope of individual liberty, not with the content of governmental language.⁴³¹

Yet a third idea available in the construction of M^* is the idea of deliberative democracy.⁴³² Hypothetical contractualists generally focus on governmental outcomes, not on the processes by which those outcomes are reached.⁴³³ But it might plausibly be argued that M^* governs processes as well as outcomes—that political institutions are required by M^* to “instantiate the ideal” of “democratic deliberation in which participants are free and equal,”⁴³⁴ quite apart from the efficacy of those institutions in assuring fair distribution, the protection of lib-

⁴²⁸ Paul J. Weithman, *Contractualist Liberalism and Deliberative Democracy*, 24 PHIL. & PUB. AFF. 314, 314 (1995).

⁴²⁹ See D'AGOSTINO, *supra* note 417, at 38-55.

⁴³⁰ See T.M. SCANLON, *WHAT WE OWE TO EACH OTHER* (1998).

⁴³¹ See RAWLS, *supra* note 417, at 289-371.

⁴³² See generally DELIBERATIVE DEMOCRACY (Jon Elster ed., 1998); DELIBERATIVE DEMOCRACY: ESSAYS ON REASON AND POLITICS (James Bohman & William Rehg eds., 1997); AMY GUTMANN & DENNIS THOMPSON, *DEMOCRACY AND DISAGREEMENT* (1996).

⁴³³ See Weithman, *supra* note 428, at 315 (noting that contractualists, at least apparently, are subject to the criticism that they are committed to an “instrumental assessment of institutional possibilities for recognizing political rights and liberties: [c]ontractualists think these possibilities are to be assessed by their effectiveness at promoting ends of which contracting parties would approve”).

⁴³⁴ *Id.*

erties, or the other outcomes required by M^* . If, in turn, the procedural requirements of deliberative democracy include a requirement of *public justification*—of reason-giving—then an expressive requirement of sorts has indeed been introduced into M^* .

Plausibly, deliberative democracy does require public justification. Amy Gutmann and Dennis Thompson contend that “[t]he reasons that officials and citizens give to justify political actions, and the information necessary to assess those reasons, should be public.”⁴³⁵ They defend this claim as follows:

First, only public justifications can secure the consent of citizens
Second, making reasons public contributes to the broadening of moral and political perspectives that deliberation is supposed to encourage. . . .
Third, reasons must be public to fulfill the potential for mutual respect that deliberation seeks by clarifying the nature of moral disagreement.
Finally, the self-correcting character of deliberation—its capacity to encourage citizens and officials to change their minds—would be undermined if reasons for policies could not be openly discussed.⁴³⁶

But this expressive requirement (like the veracity requirement mentioned above) is quite different from the purported requirements hitherto described by expressivists.⁴³⁷ First, like the veracity requirement, the justification requirement is descriptive and generic. Second, also like the veracity requirement, it presumably does not normally constrain the legal rights, duties, powers, liabilities, and other positions imposed by government, but rather applies to the descriptive sentences uttered by government along with (or instead of) the creation of legal positions. Prescriptive utterances such as “No person shall drive more than fifty miles per hour” or “Christians are exempt from generally applicable requirements that burden their religious practices” or “No person of African descent shall marry a person of a different race” do not, themselves, constitute lies or justifications.⁴³⁸

⁴³⁵ GUTMANN & THOMPSON, *supra* note 432, at 95; *see also id.* at 101 (“In a deliberative democracy, . . . the principle of publicity requires that government adopt only those policies for which officials and citizens give public justifications.”). *See generally* Frederick Schauer, *Giving Reasons*, 47 STAN. L. REV. 633 (1995) (describing the practice of reason-giving by legal officials and presenting considerations for and against that practice).

⁴³⁶ GUTMANN & THOMPSON, *supra* note 432, at 100-01.

⁴³⁷ *See supra* text accompanying notes 407-09 (noting various differences between the veracity requirement and purported requirements defended in the literature on expressivism).

⁴³⁸ Or at least not lies covered by the no-lying constraint. *See supra* text accompanying notes 410-11 (discussing the scope of the no-lying constraint). The justification requirement, like the veracity requirement, constrains the descriptive sentences ut-

Finally, the justification requirement is a *second-order* requirement. It presupposes some set of first-order moral factors applicable to governmental decisions (the first-order component of M^*), which could well be wholly nonexpressive. Whatever those considerations are, government is then under a second-order obligation to articulate them. Contrast this with, for example, the expressive theory of punishment, which purports to impose a first-order obligation on government to respond to wrongdoing with condemnation; or with the expressive theories of the Equal Protection and Establishment Clauses, which purport to impose a first-order obligation on government not to stigmatize persons or to endorse religion.⁴³⁹ The justification requirement is not expressivist in the sense of building expressive components into our collective moral framework, but rather—and more weakly—in requiring that this framework, whatever its structure, be described to the public.

CONCLUSION

In this Article, I provided a reasonably full and precise definition of an “expressive theory of law”—a definition that tracks our pretheoretical intuitions and that also conforms to the usage prevalent in much, if not all, of the scholarly literature. This definition construes expression as *linguistic*—an expressive action is one that possesses linguistic meaning, specifically by triggering a sentence-meaning convention—and an expressive theory of law as a theory that accords foundational or canonical relevance to certain types of sentence meaning possessed by the actions of legal officials. I then discussed, in considerable detail, those expressive theories that have been the best developed: the theories of punishment, of the Equal Protection Clause, of

tered by government along with (or apart from) its prescriptive sentences, and not the prescriptive sentences themselves. I think it is therefore fair to say that the justification requirement, like the veracity requirement, does not really ground an expressive theory of law. *See id.*

⁴³⁹ The expressive theory of punishment does not entail that there exist further moral requirements applicable to a government action in response to wrongdoing, beyond the requirement of condemnation. Similarly, the expressive theories of the Equal Protection and Establishment Clauses do not entail additional moral requirements beyond the no-stigma and no-endorsement rules. By contrast, a justificatory theory presupposes some further requirements, applicable to the justified action, that ought to be articulated. To say that (1) government must articulate the moral requirements relevant to its actions and that (2) there are no moral requirements for governmental actions except (1), is incoherent or nearly so. That is the sense in which the justificatory theory is “second-order,” while the other theories just described are not.

the Establishment Clause, and of regulation, each of which has been fleshed out and defended at some length by legal scholars. But each of these particular theories is, in fact, unpersuasive—or so I tried to show—and I presented a further and more general argument why expressivism is misguided. The moral factors upon which expressivists standardly rely, such as culture, self-respect, desert, or deontological norms, and which *do* figure in a sophisticated (consequentialist or deontological) moral theory, are not in fact expressive. They do not refer to linguistic meaning but, rather, to other act-properties only contingently connected to the property of possessing a given meaning.

If my arguments are correct, is there anything left for the expressivist to defend? First, and uncontroversially, the linguistic meaning of governmental action can have a *moral impact*. This is a point that I stressed in Part I, and it bears reemphasis here. Where action *A** possesses one meaning and action *A* possesses another, *A** can fare better than *A*, morally speaking, even though the underlying moral theory is nonexpressive. For example, suppose a particular activity is strongly dispreferred by most of the population (such that stopping the activity is morally optimal, given an underlying theory of preference-utilitarianism). Suppose further that *A** means “the activity is bad” and *A* means “the activity is good.” If the performance of *A** would cause the activity to cease while the performance of *A* would not, then one linguistic utterance (*A**) would have a favorable moral impact relative to a different linguistic utterance (*A*). This is true even though preference-utilitarianism is the hallmark of a nonexpressive theory.

The sheer proposition that linguistic meaning can have a moral impact, within a nonexpressive theory, is not only uncontroversial, but banal. No legal or moral scholar should be interested in defending that proposition, for no one should reasonably want to deny it. What is interesting and worth defending is the claim that particular *kinds* of meanings have particular *kinds* of impacts, within particular nonexpressive theories. There are certainly legal scholars who have tried to defend such a claim, and sometimes they describe themselves as “expressivists.” By my lights, that description is incorrect (insofar as it merely rests upon the scholar’s ascription of moral impacts to particular meaning-types); but the project (whatever we call it) of delineating such impacts surely is an important one. Nothing in this Article is meant to cast doubt upon the worth of that project.

For example, Professor Dan Kahan, in his work on criminal pun-

ishment, has argued that condemnation has a deterrent impact. A governmental action, the sentence meaning of which is condemnatory with respect to some criminal behavior, has the causal tendency to prevent would-be criminals from engaging in the behavior.

Empirical studies show that the willingness of persons to obey various laws is endogenous to their beliefs about whether others view the law as worthy of obedience: if compliance is perceived to be widespread, persons generally desire to obey; but if they believe that disobedience is rampant, their commitment to following the law diminishes. Even a strong propensity to obey the law, in other words, can be undercut by a person's "desire not to be suckered." When the law effectively expresses condemnation of wrongdoers, however, it reassures citizens that society does indeed stand behind the values that the law embodies.⁴⁴⁰

Similarly, Professor Sunstein, in his recent work on expressivism, has described the causal effect of governmental expression in shaping social norms:

[T]here is a subtler and more interesting class of cases [than where law merely supplants norms], [cases] of special importance for understanding the expressive function of law. These cases arise when the relevant law announces or signals a change in social norms *unaccompanied by much in the way of enforcement activity*. Consider, for example, laws that forbid littering and laws that require people to clean up after their dogs. In many localities such laws are rarely enforced through the criminal law, but they have an important effect in signaling appropriate behavior and in inculcating the expectation of social opprobrium and, hence, shame in those who deviate from the announced norm.⁴⁴¹

In these passages, both Sunstein and Kahan are making different versions of the generic, causal claim that governmental disapproval tends to diminish the activity disapproved (either by changing actors' beliefs, or by changing their preferences, or through some other mechanism). This causal claim presupposes nothing about the moral theory that *justifies* a reduction in the disapproved activity. Such a reduction could be warranted because the activity is harmful, or deontologically wrong, or because it is generally dispreferred, or because it decreases overall well-being (construed now in objective-list fashion), or because it effects a redistribution of primary goods from poor to rich, or from the deserving to the undeserving. The theory is, for all we know, completely nonexpressive. Therefore, the Sunstein/Kahan causal claim—and, more generally, any causal claim linking particular

⁴⁴⁰ Kahan, *supra* note 9, at 604.

⁴⁴¹ Sunstein, *On the Expressive Function of Law*, *supra* note 20, at 2032.

kinds of governmental speech to various states or events via preference-changes, or belief-changes, or in some other way—is left wholly undisturbed by the arguments presented in this Article.

Claims to the effect that governmental speech has a moral impact—that one kind of linguistic meaning is better than another or than no meaning at all, in light of the *nonexpressive* factors at hand—will typically be causal claims. I suggested in Part I that speech could have a noncausal moral impact—for example, where some person has an intrinsic preference that government utter a particular sentence-type—but that would seem to be a fairly special case. Typically, then, moral claims for and against a given kind of governmental expression will need to be evaluated in just the same way that causal claims are evaluated. The claim will be persuasive just insofar as (1) there is sufficient empirical evidence for the claimed, causal regularity; (2) there is sufficient theoretical warrant for the claimed, causal regularity; and (3) the case at hand (the case of this particular kind of speech, in this particular state of the world) is subsumed under the regularity. Arguments about the moral impact of speech, if causal, will need to be tested at the bar of physical, psychological, and social science. And if the causal argument fails that test, then it will need to be qualified or rejected. For example, the claim that governmental condemnation of a criminal behavior always deters that behavior has been empirically disconfirmed, as Kahan himself implicitly concedes:

Consider possession of guns in inner-city public schools. This behavior is infused with social [significance]. Possessing a gun confers status because it expresses confidence and a willingness to defy authority. By the same token, not possessing one signals fear, and thus invites aggression. Policies that aim at suppressing possession usually fail; indeed, when authorities aggressively seek out and punish students who possess weapons, their behavior reinforces the message of defiance associated with guns⁴⁴²

Thus, governmental condemnation is *not* universally appropriate within a deterrence theory. Rather, it will be appropriate only to the more limited extent that (based on our best understanding of the causal regularities under which governmental speech is subsumed) condemnation actually deters.⁴⁴³

⁴⁴² Kahan, *Social Influence, Social Meaning, and Deterrence*, *supra* note 15, at 363-64; see also *id.* at 381-82 (criticizing low-certainty, high-severity crime-prevention policies on similar grounds).

⁴⁴³ Sunstein provides a different example, which makes the same point.

A possible justification for [minimum wage] legislation is expressive in nature. Some people might think that government ought to make a statement to the

In short, the proper methodology for assessing governmental speech is scientific, not moral—or so I have, in effect, argued in this Article. To be sure, we need a moral framework in light of which the causal upshots of governmental speech can be evaluated. But the correct framework, whatever precisely it consists in, will not incorporate a factor that refers to the content of governmental speech. Within this framework, the question whether government ought to perform, or refrain from performing, a given utterance should be addressed in a purely scientific way. Ordinary moral reasoning—what Rawls calls the method of reflective equilibrium⁴⁴⁴—will be out of place. For example, we may intuit that a particular kind of governmental utterance is morally appropriate: condemnation may seem, intuitively, the appropriate response to crime; utterances proclaiming life or the environment to be priceless may seem the appropriate response to risky or environmentally harmful activities. And intuitive judgments are, indeed, part of the method of reflective equilibrium. But if I am right that the correct moral framework is nonexpressive, then these particular intuitions—intuitions about the appropriateness of governmental speech—will be worthless, because the method of reflective equilibrium is itself inapplicable.⁴⁴⁵ All that matters, morally, about the condemnatory or life-and-environment-valorizing utterances will be their contingent (typically, causal) connection to the further states or events that do have foundational relevance within our nonexpressive theory. It is these contingent (typically causal) connections, *nothing more*, that legal scholars who are interested in the expressive dimen-

effect that human labor is worth, at a minimum, \$X per hour; perhaps any amount less than \$X seems like an assault on human dignity. But suppose too that the consequence of the minimum wage is to increase unemployment among the most vulnerable members of society. . . . [If so] it is hard to see why people should support it.

Sunstein, *On the Expressive Function of Law*, *supra* note 20, at 2046-47. For a detailed discussion of the perverse causal consequences that the expressive content of statutes can have, in the area of environmental law, see John P. Dwyer, *The Pathology of Symbolic Legislation*, 17 *ECOLOGICAL Q.* 233 (1990).

⁴⁴⁴ See RAWLS, *supra* note 379, at 48-51 (describing this method); NORMAN DANIELS, *JUSTICE AND JUSTIFICATION: REFLECTIVE EQUILIBRIUM IN THEORY AND PRACTICE* 1-17 (1996) (same).

⁴⁴⁵ That is, once we have used the method of reflective equilibrium to arrive at a nonexpressive moral framework, the moral status of an expressive action is to be determined by looking at its status within that framework—in particular, by looking to its causal upshots—and not by giving weight to the intuition that a particular type of expression is intrinsically wrong. Given the premise that morality is nonexpressive, the intuitions are unimportant. Of course, the intuitions could prompt us to re-engage in reflective equilibrium and abandon the premise; but the very point of this Article is to argue that, at the end of reflective equilibrium, the premise survives.

sion of law should be concerned with delineating.

What about redefining "expression"? The term "expression" might be construed as denoting something other than linguistic meaning. The "expressive" cast of a governmental action would then be some property of that action other than its possession of a sentence meaning, and a genuinely "expressive" theory of law would be a theory that accorded foundational or canonical relevance to that (nonlinguistic) property.

Meaning has both linguistic and nonlinguistic variants, as discussed in Part I.⁴⁴⁶ An action nonlinguistically "means" what it provides evidence of,⁴⁴⁷ in the way that spots "mean" measles, or clouds "mean" rain, or the enactment of a spending program targeting a particular district "means" that beneficiaries are well-organized and the representative from that district fairly powerful. I am highly skeptical that basic moral factors refer to the nonlinguistic meaning of actions, particularly governmental actions. Why should the fact that action A evidences state *S* (measles, rain, the beneficiaries' organization) be itself a morally important feature of A? That feature is *epistemologically* important—given the performance of A, we now know something more about the world, namely *S*—but it is a glaring category mistake to conflate *epistemic* with *moral* importance. Spots mean measles; that hardly implies that I have moral reason to make you spotted. But this is simply the bare sketch of an argument against a genuine "expressive" theory of law, with "expressive" construed as nonlinguistic meaning. This Article does not purport to provide a full argument against such a theory; my focus has been on genuine expressive theories in the linguistic sense.

I am not aware of any legal scholar who has pursued the task of developing an "expressive" theory grounded upon nonlinguistic meaning. What some scholars have done, implicitly or explicitly, is to redefine expression in a different way—by equating the "expressive" cast of governmental action with its cultural impact. A law is "expressive," in this sense, if it shapes or reinforces social norms, social forms, or social practices more generally. Consider what Robert Cooter has to say about the expressive function of law:

⁴⁴⁶ See *supra* text accompanying notes 70-72 (distinguishing linguistic from nonlinguistic meaning).

⁴⁴⁷ More precisely, Gricean nonlinguistic meaning involves a causal connection between what is evidenced and the (nonlinguistically) meaningful actions. See Hurd, *supra* note 72, at 953 ("[Such events] function as symptoms of conditions in the world to which they are causally related."). This qualification is not important here.

A system of social norms typically has multiple equilibria. . . . A focal point can tip the system into a new equilibrium. The process of changing the equilibrium can create or destroy a social norm without changing individual values. Creating focal points is the first expressive use of law.

In addition, law can change the individual values of rational people. Internalizing a social norm is a moral commitment that attaches a psychological penalty to a forbidden act. A rational person internalizes a norm when commitment conveys an advantage relative to the original preferences and the changed preferences. . . . Changing individual values is the second expressive use of law.⁴⁴⁸

The claim here is not that expression, defined as linguistic meaning, has a causal effect on norms, forms, or practices. Rather, the claim—a conceptual or definitional one—seems to be that a law’s “expressive” property just *is* its cultural impact. On this conceptualization, a linguistically meaningful law can, but need not, be “expressive,” and a linguistically meaningless law can also be “expressive.” Other legal scholars writing about the expressive dimension of law, such as Professors Pildes and Sunstein,⁴⁴⁹ have at least intermittently made the same conceptual move as Cooter, employing “expression” as denoting a property of governmental action that is only contingently linked to language and, instead, is logically and noncontingently linked to the action’s cultural impact.

An “expressive theory of law,” thus construed, is a moral theory such that the cultural impact of a government action figures in the most perspicuous description of that action, within the theory. At various points in this Article, I have suggested that *this* kind of theory—for clarity, let us call it a “genuinely cultural theory of law”—may indeed be correct. It may indeed be the case that the correct moral theory incorporates *cultural* factors—factors that refer to social norms, forms, or practices, such that actions producing world-states identical

⁴⁴⁸ Cooter, *supra* note 43, at 586. I have omitted some language that makes Cooter’s definition of “expression” more ambiguous. At least on one plausible reading (the reading that gives less weight to this omitted language) Cooter is simply redefining expression in cultural, norm-related terms.

⁴⁴⁹ See Pildes, *supra* note 293, at 942 (“Public programs . . . *mean* something, whether this meaning is talked about in terms of their expressive character, their role in sustaining and creating a particular public culture, or the way in which understandings of public programs directly influences their implementation.”); Pildes & Sunstein, *Reinventing the Regulatory State*, *supra* note 8, at 66 (noting that “expressive dimensions” of legal and political choices “might be understood as cultural consequences of choice”); Sunstein, *supra* note 32, at 2430-32 (providing a cultural definition of stigma).

in the norms, forms, or practices that the world-states instantiate must fare the same with respect to the factors. Cultural factors may emerge within morality because well-being is both objective and partly dependent on social forms (thus Joseph Raz's theory of well-being);⁴⁵⁰ or because the primary good of self-respect is linked to social status (thus the link persistently drawn in the literature on the Equal Protection Clause and the Establishment Clause⁴⁵¹); or perhaps in some other way. Although I have not gone so far as to endorse a "genuinely cultural theory of law," I have attempted, at various junctures, to distinguish between that kind of theory and a linguistic-meaning theory, and to make clear that my arguments against an expressive (linguistic) theory of law do not necessarily apply to the expressive (cultural) variant.

But legal scholars who want to pursue the project of developing and defending a cultural theory of law need to be crystal-clear about the concepts that they deploy. If they wish to use the term "expressive" to denote that theory, then they need to stipulate that the term has been given this (fairly esoteric) referent and that the conceptual link to language has been severed. Terms like "expressive" and "social meaning" are risky and confusing; they ought to be handled very carefully. Anyone who equivocates between expression in the linguistic sense and expression in the cultural sense runs the risk of committing a fallacy of equivocation, such as the following:

- (1) "Social meaning" has intrinsic moral significance (with "social meaning" defined as the state of social norms, forms, and practices).
- (2) Law is "expressive" (in the sense of changing or reinforcing social meanings, thus defined).

Therefore,

- (3) The "expressive" properties of law (now defined as law's *linguistic* properties) are intrinsically important.

The major and minor premises of this false syllogism may well be true, but the conclusion is not. The conclusion neither follows from the premises, nor does it hold true on any other ground. (That is another way of phrasing what I have argued for in this Article.) The premises of the false syllogism bear further reflection, refinement, and perhaps eventual endorsement, but the conclusion does not.

⁴⁵⁰ See *supra* text accompanying notes 347-50 (discussing Raz's theory).

⁴⁵¹ See *supra* text accompanying notes 230-34, 281-85.

Cultural theorists, as well as others who pursue yet a further, esoteric construal of the term “expression,” would do well to keep these points clearly in mind.*

* After this Article was written, Elizabeth Anderson and Richard Pildes generously agreed to write a response. Their response is published in this volume of the *University of Pennsylvania Law Review*, along with a brief reply on my part to the response. See Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503 (2000); Matthew D. Adler, *Linguistic Meaning, Nonlinguistic “Expression,” and the Multiple Variants of Expressivism: A Response to Professors Anderson and Pildes*, 148 U. PA. L. REV. 1577 (2000). As I explain in the reply, the main thrust of the response by Professors Anderson and Pildes is not to defend the kind of “expressive theories of law” criticized in this Article—linguistic-meaning theories—but instead to present an interesting and important nonlinguistic theory. To the extent that Anderson and Pildes *do* challenge my criticisms of expressive (linguistic) accounts, I have attempted to answer their arguments in the reply rather than in this Article itself.