1997


Leslie Green
Osgoode Hall Law School of York University

Follow this and additional works at: http://digitalcommons.osgoode.yorku.ca/scholarly_works

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License.

Recommended Citation


This Book Review is brought to you for free and open access by the Faculty Scholarship at Osgoode Digital Commons. It has been accepted for inclusion in Articles & Book Chapters by an authorized administrator of Osgoode Digital Commons.
THE CONCEPT OF LAW REVISITED

Leslie Green*


Law is a social construction. It is a historically contingent feature of certain societies, one whose emergence is signaled by the rise of a systematic form of social control and elite domination. In one way it supersedes custom, in another it rests on it, for law is a system of primary social rules that direct and appraise behavior, together with secondary social rules that identify, change, and enforce the primary rules. Law may be beneficial, but only in some contexts and always at a price, at the risk of grave injustice; our appropriate attitude to it is therefore one of caution rather than celebration. Law pretends, also, to an objectivity that it does not have, for whatever judges may say, they in fact wield serious political power to create law. Not only is law therefore political, but so is legal theory — there can be no pure theory of law; concepts drawn from the law itself are inadequate to understand its nature. Legal theory is thus neither the sole preserve, nor even the natural habitat, of lawyers or law professors: it is just one part of a general social and political theory. We need such a theory, not to help decide cases or defend clients, but to understand ourselves, our culture, and our institutions, and to promote serious moral assessment of those institutions, an assessment that must always take into account the conflicting realities of life.

Those are the most important theses of the late H.L.A. Hart’s The Concept of Law, published originally in 1961. Like some other great works of philosophy, however, Hart’s book is known as much by rumor as by reading, so it will be unsurprising if, to some, that does not sound like Hart at all. For what circulates as his views — particularly, I am embarrassed to say, in law schools — is often quite different. Isn’t Hart the dreary positivist who holds that law is a matter of rules that rest on a happy social consensus? Doesn’t he think that law is objective, a matter of fact? Doesn’t Hart celebrate the rule of law and take its rise as an achievement, a mark of pro-

* Associate Professor, Osgoode Hall Law School and Department of Philosophy, York University, Toronto. B.A. 1978, Queen’s University; M.Phil. 1980, M.A., D.Phil. 1984, Oxford. — Ed. I am grateful to Denise Reaume, Jeremy Waldron, and Wil Waluchow for discussion of a number of points.
gress from "primitive" to modern society? Doesn't Hart think that liberty and justice are possible only through the certainty that clear law provides? And isn't his whole theoretical perspective straight-jacketed by a disproved, or at least outmoded, distinction between fact and value? Isn't Hart concerned more with semantics than politics?

Between those conflicting readings of — maybe I should say "attitudes toward" — Hart's book, there also lies a realm of consensus about the way The Concept of Law changed the direction of Anglo-American legal theory. For one thing, it introduced and clarified a set of questions that came to dominate the literature: Is law always coercive? What are legal rules? Do judges have discretion? Is there a necessary connection between law and morality? Hart also coined the idiom in which we debate the answers to such questions: "the practice theory of rules," "the internal and the external point of view," "primary and secondary rules," "the rule of recognition," "core and penumbra," "content-independent reasons," "social and critical morality." These terms and distinctions are now part of cultural literacy for legal theorists writing in English.

How then can there be such a wide divergence in views about Hart's theory, such confusion about his central claims? It is impossible to put it down to style. Hart is a clear and honest writer: every technical term is purchased in the coin of necessity; the occasional obscurity of language is never a cover for shallowness of thought; humor and irony he uses to lighten, not conceal. In part, it may just be that the Zeitgeist has moved on.

The Concept of Law is a book of its time. The book's language, examples, and method rest in England and, more specifically, Oxford of the fifties. Here I want to try to bridge the gap not only, as I have done in the opening paragraph, by connecting Hart's concerns with some more recent ones, but also by reexamining the 1961 work in light of some themes in its newly published Postscript.

The second edition of The Concept of Law consists of the original text together with a reply to critics that Hart left unfinished at the time of his death in 1992. The editors, Penelope Bulloch and Joseph Raz, have done an invaluable job of preparing this Post-
script for publication. Two parts were projected by Hart. The first, published here, is mainly a reply to the criticisms of Ronald Dworkin. The second, which never got beyond fragmentary notes, sought to counter other critics — Raz among them, no doubt — who, Hart concedes, found points of “incoherence and contradiction” in his work (p. 239). Hart chose to add a postscript, rather than revise the text of the book, because, as the editors note, he “did not wish to tinker with the text whose influence has been so great.”

It is not, of course, as if Hart waited thirty years to reply to his critics. He was a lively polemicist, and the points of refinement in this Postscript are less significant than a number of the essays he published after The Concept of Law. The Postscript brings no major surprises or recantations, and some of Hart’s responses to Dworkin are already well-established in the literature: there is no categorical distinction to be drawn between legal rules and principles (pp. 260-63); principles can be comprehended in the rule of recognition (pp. 265-66); judges do exercise discretion, even when they carry forward by analogical construction the underlying spirit of the law, for at some point a choice among analogies cannot be avoided (p. 275); and positivist legal theory has never been a matter of semantics (pp. 245-47). Here, I want to focus on some other points, in particular some of Hart’s last thoughts about rules, power, the connection between law and morality, and about the nature of legal theory, for there we find some of the most enduring themes, and problems, of his work.

I. LAW AS SOCIAL CONSTRUCTION

A. Antinaturalism and Antiessentialism

Constructivism is now wildly popular in the social studies, where the term has expanded to refer to almost any antirealist, antiessentialist, or antideterminist view of social life. Some of this argument is substantively idle, for it challenges no descriptive or normative thesis about its objects. If everything of interest is a social construc-

3. P. vii. It is too bad that the publisher did not share fully Hart’s view, for the new edition of this widely cited classic has inexplicably been repaginated, so that the legal theorist’s professional tools are now a copy of the first edition, together with a photocopy of the new Postscript.

4. Most of these are reprinted in his Essays In Jurisprudence And Philosophy (1983). Attention also should be drawn to Natural Rights: Bentham and Mill, Legal Duty and Obligation, and Commands and Authoritative Legal Reasons, all in his Essays On Bentham (1982); and to his brief Comment, in Issues In Contemporary Legal Philosophy (R. Gavison ed., 1987).

tion, if there is no unconstructed reality, then nothing follows from claiming that something is a social construction. Race is a social construction; and so are racism, poverty, and bullets. This might sound like a potent theory, but it is not. It is like being told that God does not exist, only to find out that the interlocutor does not believe in the existence of dogs either. Once we lose the terms of implied contrast and everything is on an ontological par, there is no critical bite to the claim.

At a lesser level of generality, constructivism sometimes simply amounts to the thesis that the object in question has a history. Here, we need to distinguish the claim that our discourse about an object has a history from the claim that the object itself does. (That the word “electron” was invented in 1890 does not suffice to show that electrons were.) The significance of constructivism about our objects of inquiry depends on whether anyone might deny the latter thesis. It is trivial to speak of the social construction of intolerance, as it is undeniably obvious that tolerance and intolerance are matters of human thought and practice. It is more interesting to speak of the social construction of race, because many people still believe that the classification of people into races is a natural one, and constructivism challenges that belief. The most potent forms of constructivism are thus those that promise to surprise us with the news that a certain object of attention owes its very existence to social history.6

Should we thrill to hear that law is a social construction? If that is just a consequence of the general thesis that everything is constructed, or that the word “law” is, then we more profitably may pass on to other business. If it is the claim that law is a phenomenon with a history, then we will have at the very least a challenge to certain arguments that associate law with reason out of time, with what P.F. Strawson once called the core of human thought that has no history.7 Some forms of ancient and medieval natural law theory might then be under threat. For example, no longer could we say, with Cicero, that

true law is right reason in agreement with Nature; it is of universal application, unchanging and everlasting . . . . We cannot be freed

6. Ian Hacking, one of the most sophisticated constructivists, puts it this way:

I respect someone who can argue that quarks are socially constructed: this is a daring and provocative thesis that makes us think. I feel a certain guarded admiration when a fact whose discovery was rewarded with a Nobel Prize for medicine is described as the social construction of a scientific fact; anyone who shares my respect and admiration for fundamental science has to sit up take notice. I do not find it similarly thrilling to read about the social construction of events that could occur only historically, only in the context of a society.

Ian Hacking, Rewriting the Soul: Multiple Personality and Sciences of Memory 67 (1995) (footnotes omitted).

from its obligations by senate or people, and we need not look outside ourselves for an expounder or interpreter of it. And there will not be different laws at Rome and at Athens, or different laws now and in the future, but one eternal and unchangeable law will be valid for all nations and all times . . . .

But even if there exists such a timeless and universal natural moral law, practically everyone agrees that human law, our law, has a history, that it is a product of human thought and practice. Naturalisms denying that are very much out of fashion among theorists — though some judges have been known to flirt with them when they run out of arguments.

Hart's theory places law firmly in history. According to him, that there is law at all follows wholly from the development of human society, a development that is intelligible to us, and the content of particular legal systems is a consequence of what people in history have said and done. Moreover, he maintains that even the normativity of law, its action-guiding and action-appraising character, is a social construction to be understood as a function of people's actions and their critical reactions to the behavior of others. For Hart, however, this is all part of the specific nature of law; it is not merely a consequence of some form of general philosophical nominalism. He theorizes law as a social construction, but it is one that emerges in a field of unconstructed reality including even certain unconstructed constraints of the human condition (p. 192).

It is tempting to see Hart's constructivism simply as a reflection of his positivism; but that would be wrong, for one of the most sophisticated positivists, and one whose influence on Hart was significant, denied parts of the constructivist thesis. Hans Kelsen held that law is a system of norms, which are not historical things at all. Kelsen did think that jurisprudence should restrict its attention to positive law — human law rather than natural law — but he did not study law under its empirical aspects. Rather, he proposed to study it as a system of norms that, according to Kelsen, exists only if it is valid. “Validity,” in turn, entails bindingness, in other words, that people ought to behave as the norms require. Historical facts — such as the fact that someone said this, or ordered that, or was disposed to behave thus — can never validate norms, for an “ought” cannot be derived from an “is.” The reason for the validity of a norm can be only another norm, and thus the ultimate reason for the validity of law must be a norm rather than a matter of fact. This “transcendental-logical presupposition” Kelsen called the Grundnorm. So while Kelsen is a positivist — law may have any


10. See id. at 201.
content and there is no necessary connection between law and morality — he is not a social constructivist. For Kelsen, law-as-norms is not really part of the social realm at all, and the tools with which we study historically situated norms — sociology, psychology, political theory, economics, etc. — are for him all “alien elements” that lead only to the “adulteration” of a pure theory of law.\(^\text{11}\)

For Hart, in contrast, law is a social construction in two senses of the term. First, law has a history. It is an institution that did not always exist, that emerged for special reasons, and that takes the form it does, including its normative character, only as a result of human action. In fact, law is a social construction of social constructions, not of brute facts but of institutional facts — namely, rules comprised by social practice and enforced by social pressure.\(^\text{12}\)

Second, Hart’s resulting concept of law is antiessentialist: Though there are central cases of legal systems, and central features of those cases, there are also borderline cases and analogical cases when without impropriety we still may speak of law.\(^\text{13}\) For Hart, there is no essence to the phenomena we call “law” and, although legal theory is right to strive to understand law’s central features, knowing these will not give us the key to all the sound generalizations about legal systems. There is no essence of law, the understanding of which can replace the hard historical, sociological, or, if you like, genealogical task of explaining law as a social phenomenon. Thus, although the term would be foreign to him, there is no coherent way for a legal theory to be more constructivist than Hart’s.

**B. Law and Social Rules**

One familiar hesitation about social constructivism and its attendant suggestion of constructing or building a reality is that it sounds all too deliberate or voluntaristic. Gender is a social construction, some even say a performance, yet people sometimes feel bullied and constrained by gender roles, and these roles often seem to resist revision. Giving adequate weight to both agency and structure, to the willed and the unwilled, was one of Hart’s major tasks in replying to the classical positivists.

---

11. See id. at 1.


13. Hart writes that:

The uncritical belief that if a general term (e.g. “law,” “state,” “nation,” “crime,” “good,” “just”) is correctly used, then the range of instances to which it is applied must all share “common qualities” has been the source of much confusion. Much time and ingenuity has been wasted in jurisprudence in the vain attempt to discover, for the purposes of definition, the common qualities which are, on this view, held to be the only respectable reason for using the same word of many different things . . . .

P. 279.
Laws, said Thomas Hobbes, Jeremy Bentham, and John Austin, are expressions of will: they are the general commands of a sovereign. But as Hart saw, such a conception of law cannot explain the variety of forms of law, nor how sovereigns can be bound by their own rules, nor how law survives the death of the commander (pp. 26-78). Above all, it cannot explain the normative character of law, the fact that it purports to impose obligations on us (pp. 82-91). A person's say-so has such normative power only when that person somehow is authorized to make norms. Most sovereign bodies are, of course, legally authorized to make law, but we need to explain the laws authorizing the sovereign as much as any other law. So if there is a norm, or norms, at the root of the legal system, it cannot be a legal norm. Then what is it? Hart rejected Kelsen's theory of the unconstructed, unsocial Grundnorm and its mysteries. Instead, he argued that fundamental lawmaking power rests on a customary social rule and has the kind of normative force that such rules have. Law, Hart argued, is a union of social rules: primary rules that guide behavior by imposing duties on people, and secondary rules that provide for the identification, change, and enforcement of the primary rules (pp. 90-99). Among the secondary rules, the so-called rule of recognition has special importance. A customary practice of those whose role it is to identify and apply primary rules, the rule of recognition provides ultimate criteria of legal validity by determining which acts create law. The rule of recognition itself is neither valid nor invalid; it simply exists as a matter of social fact or it does not. But when it does exist, and when people use it as a standard for appraising behavior, then the language of validity and invalidity comes to life, and a legal system is born.

What are these social rules of which law is constructed? Hart's account, the "practice theory," holds that there is a rule among a group \( P \), whenever there is a regularity \( R \) in their behavior such that: (1) most people in \( P \) conform to \( R \); (2) lapses from conforming to \( R \) are criticized; (3) the criticism referred to in (2) is in turn regarded as justified; and (4) \( R \) is treated as a standard for the behavior of people in \( P \). Rules are present when there is a certain kind of social practice, regular behavior together with the set of attitudes that Hart calls "acceptance" of the rule, which consists of using it as a standard for one's own behavior and that of others. Rules are to be identified and understood from the "internal point of view," the point of view of one who uses the rules as a standard,

---


15. For Hart's objections see pp. 292-93, and the essays Kelsen Visited and Kelsen's Doctrine of the Unity of Law, in HART, ESSAYS IN JURISPRUDENCE AND PHILOSOPHY, supra note 4.
though of course from the "external point of view" they merely appear as behavioral regularities (pp. 88-91).

Whether we judge this account satisfactory depends on what we expect of it. What is a "theory of rules" anyway? Here are some things we might want to know about: What are rules? What is it to follow a rule? What does it mean to say that a rule "exists"? What is the relationship between rules and reasons for action? Hart's theory does not answer all of these questions, perhaps because his original ambitions were limited somewhat. He sought to distinguish rules from other regularities of behavior, in particular from habits and predictions; to explain, as a social matter, the binding force of rules; and to set out the existence conditions for social rules. His main dispute was with two forms of reductionism: the coercion-based theories of classical positivism, which conceived of rules as orders backed by threats, and the behaviorist accounts influential among legal realists, which conceived of rules as predictions of official action. Against these, Hart's arguments are decisive. So the practice theory may be judged as a set of existence conditions for social rules, as a test for the presence of rules, whether or not it offers an analysis of rules or a full account of what it is to follow a rule. A good set of existence conditions should turn up rules when practiced, and not otherwise, but it need not itself tell us everything, or even much, about the nature of rules any more than a litmus test for the presence of an acid will tell us much about what acids are.

Hart's theory has, however, met with much criticism. There seem to be rules that are not social practices (e.g., individual rules); there are social practices that are not rules (e.g., certain common openings in chess that are not among the rules of chess); and citing a valid rule is often itself meant as a justification for one's behavior, not merely a sign that there is some other acceptable justification for it. The practice theory accommodates none of this. Moreover, matters of duty and obligation — which according to Hart amount to rules with a certain content that are enforced by serious social pressure — in some cases appear not to depend on rules at all. One certainly can believe that one has an obligation to save a drowning person without believing there is a social practice of doing this, nor even that there should be such a practice.

Dworkin argued against Hart that some rules are a matter of concurrent practice when people converge for common reasons independent of their agreement, and that at best Hart's theory must


17. They are believed important and may conflict with immediate self-interest. See pp. 86-87.
be restricted to *conventional* practices, when the fact of agreement in action is essential to the rule. So, for example, it is a merely conventional practice that in the United States people normally drive on the right because it is a sufficient reason for conformity that each expects everyone else to do the same. It is, however, a concurrent practice that people refrain from torturing others; they judge it wrong for reasons independent of what others are doing. We might add another distinction: there are also *coincidental* practices that are agreement independent, when people do the same thing for different reasons.

Dworkin attempted to confine the practice theory to conventional rules and then argued that, even there, endemic controversy about the scope of rules shows that duties must have another foundation. Conventional practice, he argued, never constitutes a normatively binding rule; it is relevant only because of the way it expresses attitudes, gives rise to expectations, etc. that may figure in the justifications for such rules. Dworkin thought that he thus had proved that judicial duty cannot be limited to the scope of a practice rule.

In the Postscript, Hart accepts Dworkin's distinction and now proposes to confine his theory to conventional rules only and to abandon the rule-based explanation of all duties. He maintains, however, that the practice theory gives a good account of conventional rules, including the rule of recognition, "which is in effect a form of judicial customary rule existing only if it is accepted and practiced in the law-identifying and law-applying operations of the courts" (p. 256). There are no a priori limits, however, on the content of such a rule, and Hart suggests that Dworkin's holistic interpretive theory might be understood as "merely the specific form taken in some legal systems by a conventional rule of recognition whose existence and authority depend on its acceptance by the courts" (p. 267).

There are a number of difficulties here. For one, the distinction between conventional and concurrent reasons for conformity does not map onto a distinction between two kinds of rules, for both kinds of reasons may be present in one rule. Indeed, it is hard to think of practiced rules that have only concurrent reasons for conformity: even a prohibition on murder must, at the margins, draw certain conventional lines around the definition of murder. Hart says that a rule is conventional provided "the general conformity of a group to [it] is part of the reasons which its individual members have for acceptance" (p. 255). To be "part of" the reasons for acceptance is a weak condition that is satisfied by all the mixed cases. Indeed, in view of the fact that general conformity with law is a

---

public good, conventional reasons for conformity are normally present. For example, even people who think considerations of justice a very good reason for treating the requirements of the Internal Revenue Code as action-guiding typically conform their behavior to it only if they believe most other people will too.

In these examples, concurrent reasons are present along with conventional ones. Sometimes Hart seems to have something else in mind. For example, at one point he calls the rule of recognition "a mere conventional rule accepted by the judges and lawyers of particular legal systems" (p. 267). Is a rule a "mere" convention only if there are no concurrent reasons for compliance? It seems very unlikely that the rule of recognition is purely conventional; officials normally have moral views about the propriety of legislative power. Yet Hart says:

Certainly the rule of recognition is treated in my book as resting on a conventional form of judicial consensus. That it does so rest seems quite clear at least in English and American law for surely an English judge's reason for treating Parliament's legislation (or an American judge's reason for treating the Constitution) as a source of law having supremacy over other sources includes the fact that his judicial colleagues concur in this as their predecessors have done. [pp. 266-67] Judicial recognition may, however, include that fact as a necessary but insufficient condition, such that each judge recognizes the Constitution as supreme law only because other judges do too and also because she for her part thinks that the Constitution is just. Hart would say that this is possible, but not necessary, for acceptance of a conventional rule can rest on anything whatever. He therefore rejects not only Dworkin's claim that there must be good moral grounds for doing what the rule says, but even the weaker thesis that people must believe, rightly or wrongly, that there are good moral grounds for doing what it says.

In the case of purely conventional rules, what motivates conformity? Despite Hart's suggestion that it may be anything whatever, I think that it must rest on something like an overlapping mutual interest19 and that the content of convention, contrary to Dworkin, fulfills not a justificatory function, but the identificatory function of showing which of the possible ways of conforming is the one that will be practiced and thus serve that interest. People typically do, of course, have preferences among alternative common ways of acting. They do not regard them all as equivalent and the differences among them are often important; but if the rule is to be purely conventional, then at some point these differences must overwhelm the divergence of interest. The worry, however, is that when we come to fill out the idea of a mutual interest in conformity

we may end up without an obligation-imposing rule. It is odd to think, for instance, that there is an obligation to eat with one’s knife in the right hand, even when it is conventional to do so. In reducing the rule of recognition to a purely conventional rule, Hart ends up without an adequate account of its binding force, not even one consistent with his own theory of social obligation.

None of this has anything to do with any indeterminacy or controversiality of conventions; the problem arises even when they are quite clear. In fact, as Hart says, indeterminacy and controversiality are side issues. The idea that controversiality of some legal decisions disproves the existence of a generally accepted rule of recognition rests on a misunderstanding of the function of the rule. It assumes that the rule is meant to determine completely the legal result in particular cases, so that any legal issue arising in any case could simply be solved by mere appeal to the criteria or tests provided by the rule. But this is a misconception: the function of the rule is to determine only the general conditions which correct legal decisions must satisfy in modern systems of law. [p. 258]

That is right, though it is put somewhat loosely, for whether a legal decision is “correct” will depend on a great many things: whether the judge can read English, follow logic, reason morally, and so forth. The rule of recognition, at least as originally conceived by Hart, did not purport to set any such conditions, generally or otherwise. It purported only to identify which of various social standards are legally relevant — which are sources of law. Because these are sources that judges are legally bound to apply, however, any analysis of the fundamental rules of a legal system still must be consistent with their obligatory force. Unlike Kelsen and Dworkin, who are on this point in agreement against him, Hart is a social constructivist about obligation itself. But his view that the fundamental rules are “mere conventions” continues to sit uneasily with any notion of obligation.

II. Consensus, Domination, and Power

So law, for Hart, is socially constructed from rules, which are themselves constructed from individual practice. That may sound, to some, like a rather complacent and comfortable view of an institution that is an instrument of social control, for accepted social rules are the sort of things that we ordinarily find in the discredited “consensus” theories of the social functionalists and systems theorists. What about conflict and power?
A. Hart's Whig History

It is unfortunate that Hart introduces his concept of law through a somewhat wooden, fictional history of social development, tracking the change from what he sometimes calls a "primitive" form of community based solely on rules of obligation to a different form of social organization based on primary and secondary rules (pp. 91-99). It is perhaps these passages more than any others that have contributed to his undeserved reputation as an enthusiast of the rule of law, for on a casual reading it seems like a Whig history of progress, from primitive to modern, from custom to law. On this reading, we begin with a primitive society in which social order rests on a broad consensus about the so-called primary rules of obligation that are maintained by diffuse social pressure. But these societies are static, inefficient, and fraught with uncertainty. Law emerges to — maybe even emerges in order to — cure these defects. Thus in a "developed," "complex" society things are better. We gain certainty, dynamism, and efficiency through the introduction of rules of recognition, change, and adjudication. That is a common reading, but it is mistaken, and the mistake easily can corrode one's understanding of Hart's theory. So let us take it more slowly.

For Hart "primitive" here just means simple. The reason that such societies do not have law is that they do not need it: nothing in human nature or society requires that we have law; many people have gotten along well without it (p. 91). A legal system, with its institutionalized means of social control, is "not a necessity, but a luxury" (p. 235). What then does Hart mean when he speaks of the "defects" that the secondary rules cure? It is crucial to his argument that simpler forms of social order do work, but only in certain contexts. He writes: "only a small community closely knit by ties of kinship, common sentiment, and belief, and placed in a stable environment, could live successfully by such a regime of unofficial rules" (p. 92). Despite this real viability, "[i]n any other conditions such a simple form of social control must prove defective and will require supplementation in different ways" (p. 92; emphasis added). It follows, then, that the "defects" that law "remedies" are not defects in simple, transparent forms of social order. They are, in a sense, defects in us. These problems arise when we try to apply those informal solutions native to a world of transparent solidarity to our world, a world of few or repudiated kinship ties, in which sentiments are not shared, in which the environment is a maelstrom of change, where, as Marx says, "all that is solid melts into air."20 So Hart’s targets here are not simple, transparent forms of social

order. It is the mistaken drive to apply techniques of governance appropriate to those targets to a more opaque world of strangers — the mistake of a Rousseau or of some modern communitarian political theory. It is also the mistake, I think, of Dworkin when he suggests that political obligation may rest on fraternal “obligations of community.”

It needs to be said, however, that Hart’s discussion in these pages is not exactly paralyzed with rigor. As many commentators have noticed, his classification of rules as primary and secondary is deployed here in a number of inconsistent ways: to mark a distinction of social importance (necessary vs. optional), of genesis (first vs. later), of normative type (duty-imposing vs. power-conferring), and of object (rules about behavior vs. rules about rules). These distinctions are neither the same nor extensionally equivalent — they will carve up the territory in different ways. Moreover, it seems unlikely there have existed societies that had rules of obligation, yet had no ways to create, extinguish, or vary such obligations. Hart focuses on a simple regime that regulates “free use of violence, theft, and deception” (p. 91), but surely the human condition also requires that any society find some way to regulate property and kinship. That suggests that well before the emergence of the systematizing rules of a legal order secondary rules in some of Hart’s senses already would have existed.

We can remedy this defect easily enough, and when we do so Hart’s theory clearly carries no significant bias in favor of modern legal orders. Paradoxically, the bias more often lies with those who detect in Hart’s argument a form of modernist triumphalism, for plainly what irks them about Hart’s story is the thought that, if a society lacks a legal system, then it lacks one of the achievements of modernity — that it is uncivilized. Because it would be parochial and demeaning to think “primitive” societies uncivilized, by *modus tollens*, they must have legal systems. Hart’s account, in failing to acknowledge these as legal systems, therefore improperly must favor modern or Western law. The logical form of this argument is a bit startling, as it attempts to deduce an “is” from an “ought,” but quite apart from that it simply relies on a premise that Hart strenuously rejects: Law is *not* a mark of civility or justice, or anything of the kind; it is just one way in which a complex society copes when

brilliant discussion of this theme, see MARSHALL BERMAN, ALL THAT IS SOLID MELTS INTO AIR (1982).

21. See RONALD DWORKIN, LAW’S EMPIRE 190-216 (1986). I have criticized Dworkin’s attempt to derive the duty to obey the law from communal obligations in *Associate Obligations and the State, in Law and the Community: The End Of Individualism?* 93-118 (A. Hutchinson & L. Green eds., 1989).

the direct, transparent form of social order no longer works very well.

B. Elite Domination

Hart argues against John Austin, who had a top-down, pyramidal view of law as the orders of a sovereign generally obeyed and backed up by threats.\textsuperscript{23} It commonly is acknowledged that Austin's was a crude theory, but some feel that it was at least realistic and that Hart, while making positivism more subtle, also loses its punch. For now law must rest, ultimately, on a form of social consensus, and that idea obscures the ways in which law in fact is rooted in social power.

Perhaps the most misunderstood part of Hart's view is the way in which law is, and is not, related to social consensus. Dworkin persistently has maintained, for example, that Hart cannot properly explain the depth and nature of controversy about law.\textsuperscript{24} For Hart, law may be controversial because it is a matter of open-textured social rules that, though they have a "core" of settled meaning, also have "penumbral" areas of doubt in which the applicability of the rule is, as a matter of practice, indeterminate (pp. 124-54). This is as true of the rule of recognition as it is of any rule identified by it. It too has a penumbral area of discretionary judgment, unregulated by law, but where decisions still may be appraised as better and worse by other relevant standards, including those of critical morality.

How much consensus does law in fact require? Many people would accept something like Dworkin's précis of Hart's theory:

[T]he true grounds of law lie in the acceptance by the community as a whole of a fundamental master rule (he calls this a "rule of recognition"). . . . For Austin the proposition that the speed limit in California is 55 is true just because the legislators who enacted that rule happen to be in control there; for Hart it is true because the people of California have accepted, and continue to accept, the scheme of authority in the state and national constitutions.\textsuperscript{25}

The problem with that as a theory is obvious enough. First, it is a fantasy: many people in California have no idea what the "scheme of authority in the state and national constitutions" amounts to. Some are not even aware that there is a state constitution, so the sense in which the community "accepts" it is pretty attenuated. Second, and worse, the "people of California" is an abstraction, or a legal concept itself: the irrelevance of the attitudes of those living in Tijuana is something that needs to be explained,

\textsuperscript{23} See Austin, supra note 14.

\textsuperscript{24} See Dworkin, supra note 21, at 3-11, 37-43; Dworkin, supra note 16, at 31-45.

\textsuperscript{25} Dworkin, supra note 21, at 34.
not assumed, by a theory of law. Beyond these two problems with the theory suggested in that passage, however, there is also something wrong with it as an exposition of Hart, something worse than mere imprecision.

Hart knew that the United Kingdom, for example, has power to make law for Northern Ireland. He also understood that it would be wrong to characterize this as a situation in which that latter community accepts the scheme of constitutional authority, for it is notorious that a large minority in Northern Ireland neither accepts, nor acquiesces in, that authority. Indeed, if we take the notion of acceptance at all seriously it must be the case that general acceptance of law's authority is pretty rare. Yet far from being an objection to Hart's theory, this is an entailment of it.

The idea that the social foundations of law rest on "acceptance by the community as a whole of a fundamental master rule" seriously distorts Hart's view. According to Hart, a regime of general social consensus is what precedes a legal regime. Law has a complex relation to conventional, customary rules, partly recognizing them, partly replacing them. The existence of the rule of recognition explains why there is law in California, in Northern Ireland, or in Quebec—a functioning legal system even though there is in those jurisdictions no broad social consensus on the overall scheme of authority. Hart writes:

In the simpler structure, since there are no officials, the rules must be widely accepted as setting critical standards for the behavior of the group. If, there, the internal point of view is not widely disseminated there could not logically be any rules. But where there is a union of primary and secondary rules, which is, as we have argued, the most fruitful way of regarding a legal system, the acceptance of rules as common standards for the group may be split off from the relatively passive matter of the ordinary individual acquiescing in the rules by obeying them for his part alone. In an extreme case the internal point of view with its characteristic normative use of legal language ('This is a valid rule') might be confined to the official world. In this more complex system, only officials might accept and use the system's criteria of legal validity. The society in which this was so might be deplorably sheeplike; the sheep might end in the slaughter-house. But there is little reason for thinking that it could not exist or for denying it the title of a legal system. [p. 117; emphasis added]

This passage illuminates one of the central ways that power relations emerge with law. Both social morality and custom, Hart notes, are immune to deliberate change; they evolve only gradually (p. 175). We almost might say that the emergence of law signals that a society has acquired a new capacity deliberately to control its common life, that, in a certain way, the community has come to self-consciousness. Customs and norms now can be changed forthwith, by the say-so of the rulers, by majority vote, or whatever. The
crucial fact is therefore institutionalization: the emergence of specialized organs with power to identify, alter, and enforce the social rules. This “advance,” however, brings both gains and costs: “[t]he gains are those of adaptability to change, certainty, and efficiency, and these are immense; the cost is the risk that the centrally organized power may well be used for the oppression of numbers with whose support it can dispense, in a way that the simpler regime of primary rules could not” (p. 202). These risks are real and sometimes have materialized. So, for Hart, the only consensus necessary for law is a consensus of elites; that is a direct and potent consequence of the fact that law is an institutionalized normative system.

C. Coercion and Power

While the institutionalization of law creates elites who may come to dominate society, Hart also argues that law is not, of its nature, coercive, at least not in the sense that all laws somehow amount to coercive threats. Hart denies that it is possible to reduce all power-conferring rules to sanction-prescribing rules (pp. 26-49). One relatively unsophisticated form of such reductionism goes like this: Rules empowering people to do things — e.g., to legislate, to contract, to make wills, to marry, to amend a constitution — specify ways in which these things must be done in order for the power to be exercised validly. Failure to conform in the relevant ways means that the purported exercise of power fails: the action in question is a mere nullity; it lacks the legal effect it purports to have. But is this nullity not essentially the same as the sanctions imposed by criminal law? After all, nullity can be as inconvenient, distressing, and expensive as some penalties.

Hart decisively campaigns against such reasoning (pp. 34-35). First, the undesirability of nullity is purely contingent; it may sometimes be a benefit to find that, for example, a certain contract is void. Second, and more important, the reductionist account falsifies the nature of power-conferring rules. In the case of a duty-imposing rule we can distinguish two different elements: the required standard of behavior (e.g., refrain from assault) and the reinforcing sanction (e.g., or else pay a fine or go to jail). These elements, however, are not similarly present in the case of a power-conferring rule. For example, promises that are neither under seal nor given for consideration are not something to be abstained from. In so defining contractual powers, the law does not have it in mind that people should either make valid contracts or refrain from promising. The provision failure to comply with which results in nullity is not a separable part of the rule; it constitutes the rule itself. To the Kelsenian suggestion that they are fragments of rules
that direct the courts to apply sanctions, Hart replies that this too distorts the way law really works:

It is of course very important, if we are to understand the law, to see how the courts administer it when they come to apply its sanctions. But this should not lead us to think that all there is to understand is what happens in courts. The principal functions of the law as a means of social control are not to be seen in private litigation or prosecutions, which represent vital but still ancillary provisions for the failures of the system. It is to be seen in the diverse ways in which the law is used to control, to guide, and to plan life out of court. [p. 40]

This passage nicely illustrates Hart's contextual approach to legal theory. There is no metaphysical answer to the question of the individuation of laws. To understand the law is just to explicate how power-conferring rules are used by those who use them: "Such power-conferring rules are thought of, spoken of, and used in social life differently from rules which impose duties, and they are valued for different reasons. What other tests for difference in character could there be?" (p. 41).

Those are certainly relevant considerations. But not only are power-conferring rules valued for special reasons, sometimes they also are feared and resented for special reasons. For Austin and Kelsen, reductionism springs from the assumption that to have a duty or obligation is to be subject, directly or indirectly, to coercion. The salience of coercion in their theories is underwritten by the significance it plays in our lives: we normally seek to avoid it. To this, Hart replies that coercion is real but secondary (pp. 199-200). First, it is needed only when law fails in its primary task of giving standards for guiding and appraising behavior. Second, it is conceivable that a legal system might not need coercive sanctions at all; their presence in all actual legal systems results not from the nature of law but from the exigencies of human nature.

There is, I think, yet another reason for attending to the coercive character of law: it is the most dramatic way in which law exercises power over people. Some, though not all, laws are coercive; but many more laws, though not coercive, are forms of social power. That is true even of those power-conferring rules that Hart generally refers to in a positive-sounding way as providing "facilities." Consider, for instance, the legal power to marry. It is granted subject to conditions that, even in so-called liberal regimes, normally include the requirement that the union be concluded between people of different sexes. Same-sex marriages are legal nullities. Now, it would, for the sort of reasons Hart gives, be misguided to think that what is going on here is a direct or indirect form of coercion, that people are being forced into the paradigmatic heterosexual union. There is, after all, no requirement to marry and, a fortiori, no requirement to marry one of the opposite sex. In most
jurisdictions it is not even wrong to purport to marry someone of one's own sex, and no punishment is given one who purports to conduct such a marriage; yet such marriages are null. There is, I think, something important here, and one who fails to notice the specific ways in which the legal system at this point embodies and exercises power fails to notice something specific about legal regulation. Heterosexist marriage laws exercise power in different ways from, say, criminal prohibitions on sodomy. The latter attempt to guide people's behavior by removing options or rendering them infeasible. The marriage laws do not do that. While they do not remove anyone's options, they do give options to some and withhold them from others, and they do so in circumstances in which many other legal and social consequences follow in train.

That law is in such ways deeply involved in social power will come as a surprise to few. Marxists, critical legal scholars, and feminists all have noticed it; Foucauldians revel in it. Few, however, have given enough thought to the ways in which the specific character of law contributes to its power. It is not enough to remind us of the facts of class, hierarchy, patriarchy, or disciplinary regimes. We also need to know how these express themselves through the different forms of law and how the legalization of, for instance, power-conferring rules affects the power distribution in society. Many other examples exist: laws create classifications, declare statuses, etc. True, these cannot be reduced to the coin of coercion, but one of the reasons for worrying about coercion, that it is a form of power, extends also to cases when power is exercised noncoercively. Thus classification systems and so forth may be productive uses of power: they create subjects, kinds of people, who then can be regulated in other ways. Of course, the informal social order does this too, but when productive power becomes imbued with the authority of law it raises the stakes enormously.

Hart's view about law and power was not at all rosy. He wrote: "So long as human beings can gain sufficient cooperation from some to enable them to dominate others, they will use the forms of law as one of their instruments" (p. 210). But the way the forms of law interact with social power, both repressively and productively, is something on which we need a good deal more work.

26. Among the Marxists, only Pashukanis clearly saw that we need an account of law's specificity as a means of social control.

If... we forgo an analysis of the fundamental juridical concepts, all we get is a theory which explains the emergence of legal regulation from the material needs of society, and thus provides an explanation of the fact that legal norms conform to the material interests of particular social classes. Yet legal regulation itself has still not been analyzed as a form... 

III. LAW AND MORALITY

A. Soft Positivism

One important point of clarification in the Postscript is that Hart now explicitly says that he is a "soft positivist":27 he thinks that while there may be some legal systems in which the fundamental test for law is wholly a matter of social fact, there are others in which it requires moral judgment (p. 250). He only denies that moral judgments are always required to know the law. That, of course, is to deny Dworkin's central thesis, that the law is whatever is entailed by the moral and political theory that both fits and best justifies the legal institutions in question.28

In response to Dworkin's objections to the claim that the rule of recognition is a matter of "pedigree,"29 Hart says that he never intended to limit its content in this way, that it may go beyond the mode of creation or adoption of rules and include also among its criteria both matters of fact not properly thought of as pedigree (e.g., the substantive constraints of the Sixteenth Amendment to the U.S. Constitution) and matters not of fact but of moral value (p. 250). Indeed, he says that this was always his view, and refers the reader both to the text of *The Concept of Law* and to an earlier article in which he claims to "state... that in some systems of law, as in the United States, the ultimate criteria of legal validity might explicitly incorporate besides pedigree, principles of justice or substantive moral values, and these may form the content of legal constitutional constraints" (p. 247). Acceptance by the courts and its reflection in practice is always a necessary element of the rule of recognition, but what they accept may include moral principles.

The passage that Hart cites in defense of this interpretation of his earlier work is, however, a reply to the Austinian view that a supreme lawmaking authority necessarily must be absolute, that it can make and unmake any law whatever, and that thus there is always an unlimited sovereign behind a legally limited legislature. Hart draws his important distinction between the presence of an enforceable duty not to legislate in a certain way — arguably inconsistent with supreme authority — and the absence of a legal power to legislate in a certain way. Yet this reference is somewhat puzzling if it is supposed to demonstrate Hart's allegiance to soft positivism. On Austin's theory, the sovereign is not to be identified

28. See DWORKIN, supra note 21, at 87-101.
29. See DWORKIN, supra note 16, at 17.
with the legislature but with whatever body enjoys habitual obedience while not rendering similar obedience to anyone else. Similarly, on Hart's theory the rule of recognition is not to be identified with the *constitution* but with the practices of recognition that are expressed when the constitution is applied. For whether a written constitution is a source of law is also a question for whose answer we must turn to the rule of recognition.

Hart describes the rule of recognition as a conventional judicial rule that identifies certain things as sources of law. But Hart's own analysis of cases in which a statute is certified as law by the rule of recognition in fact does not suggest that everything that is required for a proper application of the statute counts as law. He considers a case in which the legislature requires an industry to charge only a "fair rate" for its services (pp. 131-32). Although there will be some extreme cases of unfairness that clearly are proscribed by the legislation, there will be many more debatable cases that it would be both impossible and unwise to try to specify in advance. Hart says:

> The anticipatable combinations of relevant factors are few, and this entails a relative indeterminacy in our initial aim of a fair rate... and a need for further official choice. In these cases it is clear that the rule-making authority must exercise a discretion, and there is no possibility of treating the question raised by the various cases as if there were one uniquely correct answer to be found, as distinct from an answer which is a reasonable compromise between many conflicting interests. [pp. 131-32]

Law therefore always must strike some compromise between the two aims of settling things clearly in advance and leaving room for later choices in view of the relevant values and contextual facts. Indeterminacy in law is thus not only an ineliminable feature of language, it is also a desirable element of flexibility.

It is interesting to compare this view about statutory interpretation with what Hart says, above, about constitutions. Hart's analysis is that when a statute leaves it open to an adjudicator to determine what is "fair," "safe," "reasonable," and so on, it confers a certain *discretion* — a choice undetermined by law but open to reasoned justification. Yet in his new analysis of constitutional provisions referring to "due process" and "freedom of speech," among others, he imputes the chosen interpretation of those values back into the law itself, rather than, as his statutory case might suggest, to the discretionary power of the court. This is Hart's soft positivism.

In reality, the text of *The Concept of Law* is indecisive as between soft positivism and any harder version. It is understandable why Hart felt the urge to find in his book precedent for the views about law and morality that he seemed to hold when he wrote its Postscript. The fact of the matter, however, is that the distinction between stronger and weaker versions of the positivist thesis was
unknown when he wrote the book and was therefore not part of the polemical context of the time. Properly concerned about the natural lawyers on his right flank, Hart did not anticipate the forces also gathering on his left.

The resultant ambiguity means that those who endorse stronger versions of the positivist thesis also can find much of what they want in *The Concept of Law*. Resisting Hart’s last suggestion, they can draw from his original text some of the main elements necessary to a unified view of the examples of statutory and constitutional interpretation discussed above. For example, in discussing the idea of the “sources” of law, Hart comments as follows on the somewhat blurry distinction between the formal or legal sources of law, the reason why something counts as valid law, and its material or historical sources, the cause of its existence:

Where [a judge] considers that no statute or other formal source of law determines the case before him, he may base his decision on e.g. a text of the Digest, or the writings of a French jurist. . . . The legal system does not require him to use these sources, but it is accepted as perfectly proper that he should do so. They are therefore more than merely historical or causal influences since such writings are recognized as “good reasons” for decisions. Perhaps we might speak of such sources as “permissive” legal sources to distinguish them both from “mandatory” legal or formal sources such as statute and from historical or material sources. [p. 294]

This suggests that, on Hart’s account, something may be a good and proper reason for a judicial decision in an unregulated case — it even may be recognized by the courts as such — and yet not be the law precisely because it is not mandated by a binding source. Here, of course, Hart presents such a reason as a different “kind” of source: a “permissive” one. But all we need for a unified account is the idea that there may be mandatory and permissive aspects to a single source, and that the law ends where mandatory direction runs out. And that would give us strong positivism.

Hart’s resistance to this, and his attraction to soft positivism, may, I think, result from embracing a false dilemma: either there is some kind of *logical* incoherence in the idea that a rule of recognition could specify moral tests for law, or there is not. If not, then the most we can say is that such tests would be undesirable. Indeed, this sort of argument pervades Hart’s defense of a broad concept of law that includes immoral laws and legal systems over a narrow one that does not. He argues forcefully that we may and should distinguish between thinking, for example, that certain legal systems are wicked, and thinking that they are not legal systems at all (pp. 207-12). It is conceivable, though profoundly undesirable, that a legal system might ignore procedural and substantive justice. The other side of this coin is that it is conceivable, though in a dif-
ferent way also undesirable, that a legal system might provide wide-ranging tests of moral rectitude for the validity of law. That appears to be the burden of the following passage:

There is, for me, no logical restriction on the content of the rule of recognition: so far as "logic" goes it could provide explicitly or implicitly that the criteria determining validity of subordinate laws should cease to be regarded as such if the laws identified in accordance with them proved to be morally objectionable. So a constitution could include in its restrictions on the legislative power even of its supreme legislature not only conformity with due process but a completely general provision that its legal power should lapse if its enactments ever conflicted with principles of morality and justice. The objection to this extraordinary arrangement would not be "logic" but the gross indeterminacy of such criteria of legal validity. Constitutions do not invite trouble by taking this form.30

Since this arrangement is, indeed, not prohibited by "logic," it may seem that the only important questions are which indeterminacies cause trouble, and how much trouble we can tolerate. Dworkin complains that if we allow the rule of recognition to be anything other than a "more or less mechanical test" based on "matters of social history rather than matters of policy or morality that might be inherently controversial," then it cannot do what Hart says it does — namely, cure the uncertainty of a pre-legal, customary regime.31 In reply, Hart says that Dworkin "seems to . . . exaggerate both the degree of certainty which a consistent positivist must attribute to a body of legal standards and the uncertainty that will result if the criteria of legal validity include conformity with specific moral principles or values" (p. 251). In any case, he continues, "the exclusion of all uncertainty at whatever costs in other values is not a goal which I have ever envisaged for the rule of recognition" (p. 251).

These disputes about whether too much indeterminacy would flow from a rule of recognition that comprised moral principles are, however, beside the point. Some moral judgments are more certain, and are subject to broader agreement, than some factual ones. For instance, it may be more certain that it is wrong to torture innocent children for one's own amusement than it is that a legislature intended to promote the welfare of children by enacting a certain statute. The central question, and one that Hart never addresses here, is whether there might be constraints on a rule of recognition that are given neither by "logic," that is, by what is conceivable, nor by what is the most efficient mix of certainty and flexibility.

Consider, for example, the kind of authority that legal rules purport to have. In one of Hart's later papers, he characterizes this as

30. HART, ESSAYS IN JURISPRUDENCE AND PHILOSOPHY, supra note 4, at 361.
31. Ronald Dworkin, Reply, in RONALD DWORIN AND CONTEMPORARY JURISPRU-
The Concept of Law

a claim to provide reasons for acting that are both peremptory — that set aside the subject's own assessment of the merits of what is to be done — and "content-independent" — "intended to function as a reason independently of the nature or character of the actions to be done."32 This is not a point of logic. There is nothing in the meaning of the word "authority," nothing in its logical grammar, and nothing in the requirements of reason itself to show that legal authority is understood best in the way Hart suggests. That is a matter of legal and political theory, not logic. And, even if we come, as I think we should, to endorse his view of authority, nothing will follow as a matter of sheer logic about the rule of recognition. The motivation for endorsing that view, however, and most plausible accounts of why content-independent reasons find a place in practical thought, nonetheless may sit uneasily with the idea that the rule of recognition can make content-dependent considerations of morality — the force of which depends precisely on that nature or character of the actions to be done — part of the law. Indeed, this is the core of Raz's argument against soft positivism.33 It may be that those two theses can be reconciled somehow, or that one or the other should be abandoned. But Hart does neither, and that leaves his final commitment to soft positivism unstable.

B. Antifunctionalism

Inclusion in the ultimate criteria of legal validity is not the only way that morality could have a necessary connection to law. Another argument, at least as popular, rests on the purported social function of law. Lon Fuller, for instance, said that the function of law is to guide human conduct, which it can do only by conforming to certain procedural principles that he called the "inner morality" of law.34 Dworkin says that the function of law is to justify the use of coercion, and thus it must have the features necessary to do that.35 John Finnis says that the function of law is to coordinate action for the common good.36 Roger Shiner says that "the judgement of some system of norms that it is a legal system is at one and the same time a candidate judgement of critical morality that the system fulfills well some moral purpose."37 Michael Moore says: "If law is a functional kind then necessarily law serves some good

35. See Dworkin, supra note 21, at 93.
36. See John Finnis, Natural Law And Natural Rights (1980).
and thus, necessarily, law is in that way related to morality.”\textsuperscript{38} In their different ways, each of these writers finds in functionalism a link between law and morality.

Hart, while allowing that existing legal systems share a “minimum content” based on what is needed to help secure human survival (pp. 193-200), did not accept any further teleological claims about law’s inherent value. First, as he pointed out in his controversy with Fuller, the functional sense of good, as in a “good screwdriver” is not the same as the moral sense, as in “good person.” Any purposive and rule-governed human practice has certain goods internal to that practice,\textsuperscript{39} and this gives rise to a certain objectivity of language about practice-goods. It follows that just as there may be an internal morality of law in Fuller’s sense, there also may be an internal morality of murder according to which a good murder is one that scores high on the functional-excellence scale of murder. There is no shortage of nonmoral functions of law, from fairly neutral (e.g., guiding behavior) to most troublesome (e.g., elite domination). These functions all may have their internal “moralities,” but none of them establishes a necessary connection between law and moral value.

Second, even if law has social functions that are deemed, at least prima facie, morally good (e.g., maintaining order, making justice possible, etc.), there is but an oblique connection between that and the claim that a particular legal system has those values. The problem is that something may be a functional kind, and yet have minimal or even no actual capacity to perform its characteristic function. Take, for instance, a “printer driver.” A computer program is a printer driver if and only if it drives the printer. This is a nearly pure functional kind because it comprises a variety of software and can be instantiated in a variety of hardware. But what of a driver that has a bug and thus fails to drive anything? Does it cease being a printer driver? No, for we know it was designed for its function and, if fixed, still may perform it. Functional kinds need have only something like the capacity, when functioning normally, to perform their functions. So even if the ideal type of law is a valuable functional kind, this does not guarantee that every legal system will have some positive worth.

In the Postscript, Hart adds a third and final step to his argument. He denies that law is a functional kind at all. In response to Dworkin’s claim that the function of law is to license the use of coercive power, Hart writes: “I think it quite vain to seek any more


\textsuperscript{39} This idea was rediscovered by Alasdair MacIntyre. \textit{See Alasdair MacIntyre, After Virtue} 175-89 (1981).
specific purpose which law as such serves beyond providing guides to human conduct and standards of criticism of such conduct" (p. 249). The fact that law has such minimal social functions does not establish that law is a functional kind. For that stronger thesis, we must also show that the law is distinguished by its functions. This cannot be done. First, the only universal functions of legal systems are trivial abstractions, such as guiding conduct, maintaining order, and so on. Second, none of these functions is unique to legal systems. So, if law is a functional kind, it is a member of the same kind as, say, "custom" or "morality" or "religion." As Kelsen saw, all of these systems of norms may have similar ambitions or functions; they all may, for instance, prohibit murder. But they can be distinguished from each other only by their technique. Law is thus a modal kind and not a functional kind at all; it is distinguished by its means and not its end. The moral value of law depends primarily on the ends to which its means are put, and that is a contingent matter.

Hart's overall message about the relationship between law and morality is thus in one way similar to Hannah Arendt's in Eichmann in Jerusalem. What made Arendt's book so controversial was her claim that some Nazi atrocities were not singular, monstrous acts of deeply evil men, but rather the routinized, bureaucratic administration of people as if they were things. Seen from this point of view, Eichmann was the epitome of law-abidingness. What Arendt called the "banality of evil" thus may be seen also as the lawfulness of evil. That is what is so shocking — that law may be, in a phrase that Fuller derided, "an amoral datum." Some resist this idea so strongly, and are so intent to preserve the halo around the procedural virtues of law, that they are prepared to make its denial an item of faith. In his exchange with Hart, Fuller said: "I shall have to rest on the assertion of a belief that may seem naive, namely, that coherence and goodness have more affinity than coherence and evil." That is a faith that Hart, like Arendt, could not share.

IV. FACTS, VALUES, AND THEORIES

Hart's book sought, as he said in its preface, "to further the understanding of law, coercion, and morality as different but related
social phenomena” (p. v). He predicted that lawyers would see his work as an exercise in “analytical jurisprudence ... concerned with the clarification of the general framework of legal thought rather than with the criticism of law or legal policy” (p. v). He also liked to think of it as “an essay in descriptive sociology” (p. v).

In case it seems obvious that that is what a legal theory should be, it is worth remembering Dworkin’s fundamental disagreement. According to Dworkin “[j]urisprudence is the general part of adjudication, silent prologue to any decision at law.” At the opening of Law’s Empire the chapter title puts the question “What is law?”, the subtitle asks “Why it Matters,” and the first sentence answers: “It matters how judges decide cases.” On this point Dworkin never wavers. That legal theory is about adjudication is not a matter on which he ever had second thoughts. As early as 1965, he wrote: “What, in general, is a good reason for a decision by a court of law? This is the question of jurisprudence ....” Hart, the antiessentialist, never thought that anything properly could be called “the question of jurisprudence.” The first chapter of The Concept of Law is entitled “Persistent Questions,” in the plural, of which he identifies three as underlying much of the tradition of argument about the nature of law: How is law related to coercive threats? What does the obligatory force of law amount to, and how is it related to moral obligation? What are social rules and in what way is law about rules? (pp. 6-13). None of these is the question of jurisprudence, and, interestingly, none of them is Dworkin’s question.

In the Postscript, Hart aims for a peaceful coexistence: Dworkin’s question is certainly an important one, and legal reasoning and the theory of adjudication were topics on which Hart came to think he should have written more. But he insists that there remains a role for what he calls a “general and descriptive” legal theory (pp. 239-40). It is general in that it is not tied to any particular legal system or culture, and descriptive “in that it is morally neutral and has no justificatory aims” (p. 240). Hart adds that “it does not seek to justify or commend on moral or other grounds the forms and structures which appear in my general account of law, though a clear understanding of these is, I think, an important preliminary to any useful moral criticism of law” (p. 240).

44. Dworkin, supra note 21, at 90.
45. Id. at I.
47. For some influential positivist accounts of adjudication and legal reasoning, see especially, Neil MacCormick, Legal Reasoning And Legal Theory (1978), and Raz, supra note 33, at 180-209.
Is moral neutrality possible in these matters? Here, I think we are often hampered by a rather sparse vocabulary. Having abandoned a crude distinction between fact and value, it is now commonplace for legal theorists to infer that purported statements of fact necessarily presuppose or express moral and political positions. So while we have rejected what Dworkin calls a “flat distinction between description and evaluation,”\(^48\) we sometimes fall into a casual identification of the evaluative, the moral, and the political.

We should avoid this, first by observing the difference between a description and a statement of a fact. A statement of fact may be appraised as true or false. A description normally is not thought of as being true or false, but as being helpful or unhelpful, illuminating or unilluminating. There is an infinite number of possible descriptions of any object or state of affairs because there are infinitely many facts about each. A description of something is thus never a statement of all the facts about it; it is a selection of those facts that are taken to be for some purposes important, salient, relevant, interesting, and so on.\(^49\) This is not to say that a description is an appraisal of its object; it is to say that describing is always done from the point of view of certain values and in that way expresses those values.

Second, we need to remember that not all values are moral values. For example, there are the theoretical values of simplicity, consistency, fecundity, and so on.\(^50\) There is also a deep evaluative substratum to practical thought that begins with reflection about what is really most salient about the human condition (e.g., the fact that we can reason or feel pain), and what, though true, is marginal (e.g., that we are featherless and bipedal). Here begins the realm of practical value.

An illuminating descriptive account of law will, therefore, implicate values in these two ways. We will be inclined to endorse a theory that is as simple as its subject matter permits, that is consistent with most of the other views we endorse, and that produces interesting new hypotheses about and deeper understandings of law and other related phenomena. Furthermore, because law is part of human thought and practice, we also will prefer to describe it in an anthropocentric way, as it relates to those things we take to be most important about ourselves — the way law embodies power relations that can harm or help people, for instance, rather than its connection to the demand for pulp and paper. In these ways, a general legal theory must have evaluative aspects, but this stops well short

\(^{48}\) RONALD DWORIN, A MATTER OF PRINCIPLE 148 (1985).

\(^{49}\) See AMARTYA SEN, CHOICE, WELFARE AND MEASUREMENT 432-49 (1982).

\(^{50}\) Waluchow calls these "meta-theoretical-evaluative" considerations. See WALUCHOW, supra note 27, at 19-30.
of the basic features of moral evaluation on any plausible account. A moral theory will, of course, strive for some similar theoretical virtues, and any humanistic morality that it systematizes will also begin from some set of salient facts about the human condition; however, the characteristic features of moral judgments — identifying basic goods, expressing approval and disapproval, endorsing universal prescriptions, among others — all involve commitments well beyond those of description. Thus, while descriptions are not value-neutral, they need not be morally fraught either.

This is all that Hart needs. Provided that one may describe a state of affairs without thereby endorsing it, the necessary sort of neutrality is preserved. That is how we should understand his claim that "[d]escription may still be a description, even when what is described is an evaluation" (p. 244). None of the familiar arguments about the theory-ladenness of factual statements, or the value-ladenness of descriptions, undermines this idea.

Hart's most powerful argument for a general and descriptive theory of law, however, is not the soft-positivist one that holds that a Dworkinian approach is consistent with his own, but rather the potent claim that Dworkin's theory actually requires something very like Hart's for its success. Even if we say that moral principles are part of the law because they are entailed by the theory that best fits and justifies the legal institutions as a whole, we still need some way to fix on the relevant institutions in the first place. We can interpret only an object that we can identify. For this reason, Dworkin allows that there must be something called "preinterpretive" law:

[T]here must be a "preinterpretive" stage in which the rules and standards taken to provide the tentative content are identified. . . . I enclose "preinterpretive" in quotes because some kind of interpretation is necessary even at this stage. Social rules do not carry identifying labels. But a very great degree of consensus is needed — perhaps an interpretive community is usefully defined as requiring consensus at this stage — and we may therefore abstract from this stage in our analysis by presupposing that the classifications it yields are treated as given in day-to-day reflection and argument.51

Such preinterpretive agreement is, of course, contingent, local, and open to challenge and change. There must be, however, "enough initial agreement about what practices are legal practices."52 Fortunately, there is; the boundaries are clear to any lawyer who knows the craft. Dworkin just insists that lawyers do not learn them by first sharing a view about what a legal system amounts to, or even

51. DWORKIN, supra note 21, at 65-66.
52. Id.
sharing "common criteria or ground rules" for knowing which facts about the world are legally relevant.53

These ideas need clarification, for the assertions that there are boundaries but no shared criteria, social rules but no ground rules, are obscure. If the point is merely, as Dworkin says, that neither lay nor professional understandings of law result from having a good theory in pocket, then there is no dispute. Everyone admits that the theories arrive late, but what Hart and most other writers assert, and what Dworkin here seems to deny, is that interesting things may be said about "preinterpretive" law, about the very boundaries of legal practice. Dworkin talks of the "consensus," "paradigms," and "assumptions" that form preinterpretive law in a way that suggests that he takes such nodal points of agreement to be surd, unstructured, facts. Hart, in contrast, thinks they are structured and constructed by conventional social rules. Even if Hart is wrong about the character of these rules, I think it premature to conclude that there is no work for theory to do at this level. Can it really be that a consensus of judgment defies explanation? I am inclined to say no.

It is, however, a fair question to ask, as students will: "What is the use of such a theory at such a level of generality?" It is no response to say: "Well, law is interesting, and we might as well have a general theory about it." We do need to explain why these deepest, structural questions about our institutions are interesting and important. Dworkin's answer, that it matters how judges decide cases, will not help at this level, for little in a general theory of law tells us how to decide cases. Nor should we, I think, try to hitch the study to some argument about the value of communing with the Great Books of legal theory. If a general theory of law is of no real use, then we might well question the greatness of these books and direct students to any of the competing bibliographies: Brontë is always a better read than Bentham. In any case, piety about books was not Hart's style.

He argued with his predecessors, but refused to encumber the text with citations. He relegated some modest number of references to the back of the book with instructions that they be read, if at all, later. (The book would now fail as a tenure piece.) In the preface, a little diffidently, he justified this decision by referring to his pedagogical aim:

I hope that this arrangement may discourage the belief that a book on legal theory is primarily a book from which one learns what other books contain. So long as this belief is held by those who write, little progress will be made in the subject; and so long as it is held by those

53. See id. at 91.
who read, the educational value of the subject must remain very small. [p. vi]

Legal theory has a number of legitimate ambitions, and no one need pursue all of them simultaneously, or even at all. I am sure that the division of labor is part of what has provoked misunderstanding of Hart’s views. Another part is the way this division plays out in the structure of legal education at least in North America. If you come to law school with political curiosity but no prior training in the humanities or social sciences, and are then fed the typical first-year diet of doctrine, you are bound to arrive at your first legal theory course aching with hunger. To long for a first chance to debate liberty, justice, or equality and then to be served up the practice theory of rules must be pretty frustrating. All the talk about general theory being the necessary preliminary to evaluation may just sound like more excuse for delay.

There is some justice in this reaction, for description is not exactly a preliminary to evaluation. That view is influenced too much by the old idea of philosophy as underlaborer, clearing away the muddles. Description and evaluation intertwine and, ideally, cooperate. Consider Hart’s advocacy of a wide concept of law, one including as full-blooded laws those that are immoral. This is not, he insists, a matter of semantics; we should prefer one concept to the other only on grounds of theoretical fecundity or of usefulness in practical judgements (p. 209). But what he says about theoretical fecundity is brief and question-begging. He thinks it confusing to separate out iniquitous laws from the rest as “non-laws” because they have so much structurally in common with laws, and because no other discipline — such as legal history — has found it profitable to distinguish them so (pp. 209-10). This argument, however, clearly assumes that the structural features of law are the most central, and that is what is at issue. Nor can we say the broad concept is strictly necessary for clear practical deliberation. It is true that matters of political obligation, punishment, and the rule of law raise complex moral issues. If the rule of law is to be sacrificed in order that very great evil be punished, then so be it, is Hart’s response. The idea that we may face a choice of evils, however, can be preserved in other ways too. For example, one might say that there is always a prima facie moral obligation to obey the law, and also a prima facie moral obligation to do justice. The only sure route from the premise “This is the law” to the conclusion “This must be

54.
A case of retroactive punishment should not be made to look like an ordinary case of punishment for an act illegal at the time. At least it can be claimed for the simple positivist doctrine that morally iniquitous rules may still be law, that this offers no disguise for the choice between evils which, in extreme circumstances, may have to be made.

Pp. 211-12.
obeyed" is one that recognizes no countervailing considerations or that holds that the obligation to obey is absolute, but those are not errors inherent to a narrow concept of law. So I don't think that we are going to justify a general theory of law as a sufficient prophylactic to moral and political obtuseness.

The main interest in a general theory of law, I think, rests in the way that it helps us understand our institutions and, through them, our culture. It is only when we move beyond the question of how to decide cases, and even beyond the casuistry of applied ethics, that we begin to appreciate the role of a general theory. What is law that people take such pride in it? Is law a good idea? How and to whom do legal institutions distribute power? Is the rule of law always desirable? Can it help achieve justice? What might we gain, or lose, by limiting the reach of law? Those are deep and urgent questions for political theory, and also for political practice. Anyone who wants answers to them will need the help of a general theory of law. That Hart made such a good start on it is, for us, extremely lucky.
THE REAL ETHIC OF DEATH AND DYING

Norman L. Cantor*


When medical science became capable of prolonging the dying process beyond the point that most patients would wish, medical management of the dying process became a necessity. Health-care providers no longer could strive inexorably to extend waning human lives. The search thus began for an ethic to govern medical management of the dying process.¹

Peter Singer's Rethinking Life and Death,² a provocative and entertaining book, purports both to critique "the old ethic" — the book is subtitled "The Collapse of Our Traditional Ethics" — and to propound a "new ethic" to regulate the medical handling of dying patients.³ Although the book does underscore some anomalies in end-of-life care, its account of the dominant ethic of death and dying proves inaccurate. Rather than portraying the existing order — or disorder — it creates a straw man. Moreover, despite the highly problematic nature of his "new ethic," Singer defends it only superficially.

This review essay contains three parts. The first exposes the deficiencies in Singer's depiction of the old ethic. The second lays bare the key ingredients in his new ethic and discusses some of its major issues and weaknesses. The third presents my own prescription for an appropriate ethic to govern medical management of the dying process.

---


² Peter Singer is a philosophy professor at Monash University in Australia and has written extensively on animal rights and on bioethics topics such as reproductive technology and care of the dying.

³ The book touches on an ethic toward fetal life and animal life as well as dying medical patients. Its focus, however, rests on the medical handling of human life. This essay centers on that feature.
I. THE OLD ETHIC AS STRAW MAN

According to Singer, a "sanctity of life" ethic dominates the traditional approach to death and dying. A central premise of that supposed ethic is that all human beings, no matter how rudimentary their mental function and capacity, deserve protection. That protection includes a prohibition against the intentional taking of innocent human life and, in the medical context, a ban on letting patients die simply because of deteriorated quality of life (pp. 73-75). Exceptions to this sanctity-of-life approach supposedly exist to allow for the cessation of "extraordinary means of medical treatment" and for the use of analgesics that are intended to relieve pain but incidentally hasten death (p. 147). However, the strict sanctity-of-life ethic described by Singer has not prevailed in Anglo-American jurisprudence since 1976, when the New Jersey Supreme Court in In re Quinlan upheld the discontinuation of life support maintaining a permanently unconscious patient.

Singer contends that physicians who remove life-sustaining machinery with the object of allowing a patient to die take an innocent human life — a violation of what he sees as the old sanctity-of-life principle. In Singer's view, the medical profession secured authorization to take such steps in the 1993 Bland case, in which Britain's House of Lords upheld the removal of a feeding tube sustaining a permanently unconscious patient (pp. 65-66). This assertion ignores the fact that American courts for twenty years have upheld the right to remove life support, including artificial nutrition, from permanently unconscious patients even though the acting parties involved understood that death would ensue. Quinlan was the first such decision, but a succession of cases from other jurisdictions have followed suit. Singer attempts, unsuccessfully, to distinguish these American precedents as being grounded in autonomy — the prior expressions of now incompetent patients (p. 64). Quinlan did

4. Singer contends that the existing whole-brain definition of death has "come apart." See p. 36. That contention rests principally on the fact that even after brain death — under current definitions — occurs, certain endocrinal or hormonal functions of the body continue for some period. See p. 36. Yet Singer does not articulate an alternative definition of death. Instead, he reformulates the concept of "personhood" and declares that the lives of "nonpersons" merit no legal protection. I critique that reformulation of personhood, a part of Singer's "new ethic," in Part II.


7. Quinlan involved the withdrawal of a respirator, not artificial nutrition, and the patient surprisingly endured for nine years. Nonetheless, the firm medical expectation had been that withdrawal of the respirator would cause the prompt death of the patient. See Quinlan, 355 A.2d 647.

not rely on the patient's prior expressions. Furthermore, subsequent decisions have endorsed the withdrawal of life support from patients even in the absence of clear-cut prior expressions. In short, American jurisprudence on death and dying generally accepts that physicians sometimes may "take innocent life," as Singer defines the concept.

The second aspect of Singer's old sanctity-of-life ethic — the notion that poor quality of life can never justify the termination of life-sustaining medical intervention — never really has prevailed. Since 1976, American courts have recognized that a person's health may deteriorate to such a degree that she may be better off dead than alive. Cases have applied this principle to both competent and incompetent patients. For incompetent patients, judicial acceptance of end-of-life determinations has relied both on the dismal status of the patient — such as permanent unconsciousness — and on determinations that the burdens of existence, such as pain and suffering, can outweigh the benefits of extended life. Contrary to what Singer suggests, courts frequently consider diminished quality of life, in the sense of grievous bodily deterioration, in shaping the bounds of medical intervention in the dying process.

With regard to the asserted "old ethic," Singer suggests that permitting the removal of "extraordinary means" of life preservation constitutes the main deviation from a strict sanctity-of-life principle (p. 188). The concept of extraordinary means, which originated in a 1957 pronouncement of Pope Pius XII, influenced the original position of the devoutly Catholic Quinlan family. In fact, the concept sometimes was cited as a possible demarcation of permissible medical conduct in ending life-sustaining intervention. For exam-

13. In the context of assessing an incompetent patient's best interests, the quality-of-life consideration sometimes occurs under the heading of "dignity." See Norman L. Cantor, Quality of Life in Legal Perspective, in 3 ENCyclopedia of BioETHICS 1361-66 (Warren T. Reich et al. eds., 1995).
Death and Dying

ple, the American Medical Association House of Delegates used the extraordinary means terminology in 1973 in suggesting guidelines for terminal care.\textsuperscript{16}

Nevertheless, American jurisprudence long since has abandoned the ordinary-extraordinary dichotomy.\textsuperscript{17} Authorization to withhold or withdraw life support now extends to the most basic forms of medical intervention, including blood transfusions,\textsuperscript{18} artificial nutrition,\textsuperscript{19} and chemotherapy.\textsuperscript{20}

In sum, the old ethic of death and dying presented by Singer bears little resemblance to the prevailing ethic found in American cases of the past twenty years. Had Singer articulated and defended a sensible new direction in the death and dying ethic, that flaw would seem forgivable — but he did not. Although he does endorse some unconventional positions, he fails adequately to defend or even to articulate their implications. I turn to consideration of those positions.

II. \textbf{WEAKNESSES OF THE NEW ETHIC}

A. \textit{Human Nonpersons}

Singer’s new ethic centers around the notion that not all human beings are persons (pp. 180-83). To be a “person,” he says, a being must have an awareness of self over time and enough reasoning capacity to plan for the future (pp. 182, 218). Under this theory, certain human beings — including anencephalics, permanently vegetative patients, and neonates — are deemed nonpersons. On the other hand, certain nonhuman animals — including whales, dolphins, monkeys, dogs, and pigs — are deemed persons (pp. 180-82, 205-06, 209-10). Although Singer does not address it, his framework also might classify some severely retarded or demented human beings as nonpersons. This might include patients with advanced Alzheimer’s, for example.\textsuperscript{21}

Singer’s personhood framework falters in its superficial consideration of the implications for human nonpersons. Many commentators have argued that absence of neocortical function — which includes the capacity to interact with others — ought to form the


\textsuperscript{17} For a clear-cut repudiation of that dichotomy, see \textit{In re Conroy}, 486 A.2d 1209, 1234-35 (N.J. 1985); \textit{Alan Meisel, The Right to Die} 481-86 (2d ed. 1995).

\textsuperscript{18} See \textit{In re Brooks}, 205 N.E.2d 435 (Ill. 1965).

\textsuperscript{19} See \textit{In re Guardianship of L.W.}, 482 N.W.2d 60 (Wis. 1992).


\textsuperscript{21} Some of these beings have lost their sense of self over time, a factor critical to self-identity and personhood under Singer’s framework.
boundary of death. Singer does not, however, classify his nonpersons as dead. Rather, he sees them as creatures with diminished rights and expectations, retaining some interests but lacking normal protection against involuntary death (p. 198).

Singer briefly considers the implications of nonpersonhood in the context of neonates. He supports medical infanticide, at the parents' discretion, during the first few weeks of a neonate's existence, asserting that these young infants are "not yet full members of the moral community" (p. 130). In his view, the parents of a Down's syndrome neonate may withhold her life support if they prefer to raise only children better equipped to deal with life's challenges (pp. 212-15). Singer does not discuss the concomitant issues of organ harvesting, medical experimentation, or allocation of scarce medical resources; however, it seems fair to assume that his theory would favor the interests of live persons over the interests of nonperson neonates in prospective life.

Singer's approach to the implications of nonpersonhood proves even more perfunctory in the context of permanently unconscious patients. Must we honor the request of a previously competent patient to be maintained in a permanently vegetative state? Singer says that such wishes should be "taken into account," but should not be decisive (p. 192). What about the independent emotional and financial interests of the patient's relatives and other caretakers? Singer merely says that such interests "deserve consideration" (p. 192). What about the competing interests of potential organ recipients and potential beneficiaries of nontherapeutic medical experimentation on the permanently unconscious patient? While Singer comments that we "cannot ignore the needs of others" (p. 192), he does little to elucidate a hierarchy of interests regarding the treatment of human nonpersons.

Labelling permanently vegetative patients as nonpersons achieves very little. If Singer's concern is the indefinite preservation of a dismal quality of life — with no real benefit to the perma-


23. See John Harris, Euthanasia and the Value of Life, in EUTHANASIA EXAMINED 6, 19-20 (John Keown ed., 1995). Harris, who subscribes to a definition of personhood identical to Singer's, argues that the interests of nonpersons must give way to the significant needs of actual persons.
nently insensate being and with real opportunity costs to society — that concern can be met without denoting the vegetative patient a nonperson. I have argued elsewhere\(^2\) that permanently unconscious patients should be allowed to die. My rationale, however, is not that these patients are nonpersons, but that withdrawal of life support in this circumstance very probably accomplishes the result that the patients would want. The vast majority of people, when asked, say that they want no life support to maintain them in a permanently insensate state.\(^2\) We ought to respect this common, and therefore putative, wish in the absence of prior instructions or personal indications to the contrary. Furthermore, even if the vegetative patient did in fact request life-sustaining measures, this does not mean necessarily that nonpatient sources must fund this care.\(^2\)

Singer might ask in return: What do we gain by calling permanently unconscious beings “persons,” especially if we should let them die anyway? I base my response on a factor that Singer largely ignores — namely, society’s interest in sanctity-of-life, not as a mandate to prolong every human life, but as an injunction to respect the interests of human beings and humanlike beings in helpless and vulnerable states. Sanctity-of-life in that sense centers on the promotion of social sensibility to the interests of humans and of the moral tone of society.\(^2\) From this sanctity-of-life perspective, human beings ought to be deemed persons with moral status regardless of their intellectual capacities.

At the very least, personhood status should not depend upon awareness of self over time. A societal interest in moral tone compels a showing of full respect for beings with the capacity to experience human feelings and emotions.\(^2\) Calling such beings persons does not mean that we must preserve them at all costs or in situations in which their own welfare or putative preferences indicate that they should be allowed to die.\(^2\) Acknowledging personhood simply implies a respect for the significant interests of such beings.

---

24. See Cantor, supra note 8, at 410-17.


28. For views that the capacity for feelings and interactions is sufficient for moral status as a person, see Daniel Callahan, Terminating Life-Sustaining Treatment of the Demented, HASTINGS CTR. REP. NOV.-DEC. 1995, at 25; Stephen G. Post, Dementia in Our Midst: The Moral Community, 4 CAMBRIDGE Q. OF HEALTH CARE ETHICS 142 (1995).

29. For an elaboration on the appropriate criteria for allowing incompetent patients to die, see Part III.
including their autonomous choices, their human dignity,\textsuperscript{30} and their presumptive right to continued existence. It reverses Singer's ostensible indifference toward the lives and interests of those human beings, including neonates and profoundly incapacitated adults, who experience pleasure and pain despite their intellectual deficits. That indifference may stem from Singer's equating the interests of humans and nonhuman animals. In other words, his reluctance to recognize a presumptive right to life for humans with gravely diminished mental capacity may reflect an unwillingness on his part to recognize a parallel right for fish and fowl (p. 222).

\textbf{B. Transition to Active Euthanasia}

Singer cannot fathom a regime of medical management of the dying process that permits the cessation of life-sustaining medical intervention but forbids the administration of lethal poisons, or active euthanasia. For him, removal of artificial nutrition, or of any life-sustaining measures, constitutes the intentional taking of human life (p. 68). He understands that such removal is permissible in response to the wishes of a competent patient because of the patient's strong interest in shaping a dignified death and in avoiding suffering. Singer also understands that those same interests underlie any request for active euthanasia. He therefore sees a "moral incoherence" in forbidding active administration of death while permitting removal of life support (p. 80). He sees two medical actions that "are equally certain ways of bringing about the death of the patient" (p. 221). Indeed, active administration of a poison, with its immediately fatal result, seems to him more humane than a withdrawal of care, which creates a more protracted end-of-life ordeal for both the patient and her family.\textsuperscript{31} All this leads Singer to endorse, as part of his new ethic, physician-assisted suicide and active euthanasia when they fulfill a suffering patient's firm wish to die.

Singer sees it as anomalous to authorize some but not all actions that precipitate death. He mistakenly believes that the "old ethic" simply embodies an action-inaction dichotomy. In fact, that old ethic distinguishes between inaction in the face of a fatal natural affliction and the introduction of outside agents such as poisons or bullets that accelerate death. While "pulling the plug" — with-

\textsuperscript{30} "Human beings who lack or have lost the capacity for autonomous actions are nonetheless humans who retain their inherent dignity. Respect for persons comprises more than respect for autonomy." Edmund D. Pellegrino, \textit{Patient and Physician Autonomy: Conflicting Rights and Obligations in the Physician-Patient Relationship}, 10 J. CONTEMP. HEALTH L. & POLY. 47, 49 (1994).

\textsuperscript{31} Singer asks: "How can it be lawful to allow a patient to die slowly, though painlessly, over a period of weeks from lack of food but unlawful to produce his immediate death by lethal injection, thereby saving his family from yet another ordeal . . . ?" P. 78.
drawing life support, such as a respirator — is indeed an action, it traditionally has been treated, so long as it merely removes a medical obstacle to a natural death, as the moral and practical equivalent of medical nonintervention. When a dying patient objects to further medical intervention, no difference exists between failing to activate her respirator, failing to replenish her expired oxygen supply, or withdrawing her respirator. All these forms of medical behavior allow a natural dying process to run its course.

From the outset, death-and-dying jurisprudence has regarded medical withdrawal of life support as equivalent to medical nonintervention. It also has distinguished both forms of conduct — withholding and withdrawing medical intervention — from the introduction of outside lethal agents. In fact, cases upholding the prerogative of a patient to reject life-sustaining treatment uniformly have distinguished that behavior from suicide on the basis of the distinction between letting nature take its course and initiating lethal agents. Recent cases have maintained that distinction in rejecting dying patients' asserted right to physician-assisted suicide. The Michigan Supreme Court recently commented:

[W]hereas suicide involves an affirmative act to end a life, the refusal or cessation of life-sustaining medical treatment simply permits life to run its course, unencumbered by contrived intervention . . . . There is a difference between choosing a natural death summoned by uninvited illness or calamity, and deliberately seeking to terminate one's life by resorting to death-inducing measures unrelated to the natural process of dying.

The question then becomes whether a meaningful distinction exists between letting nature take its course and accelerating a natural dying process. Singer sees the difference as perverse — nonintervention, as opposed to active euthanasia, tends to prolong the dying process and to increase the burdens on patients and their caretakers.

32. See cases cited infra note 34.

33. An important policy concern reinforces the willingness of courts to treat life-support withdrawal as equivalent to noninitiation. If medical personnel cannot remove life support, they would be deterred from initiating it when a patient faces a strong chance of a protracted existence in a dismal, deteriorated state, yet has at least a slight chance of recovery. See Council on Ethical and Judicial Affairs, American Medical Assn., Decisions Near the End of Life, 267 JAMA 2229, 2231 (1992) (finding "no ethical distinction between withdrawing and withholding life-sustaining treatment"); see also Extracts from the Report of the House of Lords Select Committee on Medical Ethics, in Euthanasia Examined 96, 105 (John Keown ed., 1995).


36. Singer also perceives an anomaly in contemporary medical ethics' endorsement of analgesics that may mitigate a patient's pain but that also may accelerate her death. Again,
His position deprecates any societal interest in promoting sanctity-of-life in the sense of maintaining respect for human existence in all its forms. We promote respect for human life by limiting the circumstances in which we permit humans to terminate human life. We already tolerate war, capital punishment, and self-defense. Arguably, we should hesitate to add to this list, even for an object as appealing as the relief of suffering. Some commentators perceive a useful social message in drawing the dividing line between cessation of treatment — when disease causes the ultimate death — and euthanasia — when a human act causes the patient's demise. Cases rejecting a right of people to starve themselves to death in hunger strikes have drawn just such a line. They distinguish between a person's decision to starve and a dying person's rejection of life-sustaining treatment. Although we ultimately may come to regard the distinction between allowing and precipitating death as a shallow psychological or symbolic anachronism, it surely deserves more than the short-shrift consideration that Singer accords it.

Singer cannot fathom how society can permit active administration of possibly lethal outside agents like analgesics while proscribing active euthanasia. See p. 188. Cf. Compassion in Dying, 1996 WL 94848 (holding that a Washington statute prohibiting doctors from prescribing life-ending medication for the terminally ill who want to hasten their own deaths violates due process).

The current legal authorization of analgesic administration, at first blush, does seem inconsistent with a ban on euthanasia. The customary explanation — that physicians administer analgesics with a primary intent to relieve pain — proves unpersuasive. Physicians might commit euthanasia with the same primary intent. A better explanation lies in the difference between the criminal law's authorization of some risky yet potentially beneficial acts and its condemnation of probably lethal conduct undertaken for the same benevolent purposes. For example, a surgeon may perform an operation for an important cosmetic benefit even though the operation poses some modest risk of death to the patient, yet he cannot perform the same operation if the mortal risk is very great. Along similar lines, a physician may administer an analgesic that is necessary to relieve pain even though the act causes some risk of death. She may not, however, administer a risky dosage when a smaller dosage would do, when alternative means to relieve the patient's pain exist, or when death probably will be caused. Thus, the tension between the authorization of risky analgesics and the prohibition on euthanasia is not as great as it appears. For a full account of the legalities of analgesic administration, see George C. Thomas & Norman L. Cantor, Pain Relief, Acceleration of Death, and Criminal Law, in 6 KENNEDY INST. OF ETHICS J. (forthcoming 1996) (manuscript on file with author).


39. See also Cruzan v. Director, Missouri Dept. of Health, 497 U.S. 261, 280 (1990) (indicating in dictum that although a dying person might have a right to reject life-sustaining medical intervention, a healthy person has no comparable right to starve himself to death).

40. The distinction certainly leads to some fine line drawing. A healthy person who engages in a hunger strike initiates an unnatural dying process and therefore is regarded as committing suicide, but if a deteriorated, fatally stricken patient makes a deliberate decision to stop eating, the strong medical and legal inclination is to acquiesce in the patient's fatal...
Even if Singer is right about the shaky moral basis of the traditional line between letting die and killing, his leap to the endorsement of active euthanasia seems premature. Practical concerns cause many bioethicists to shy away from supporting the legalization of active euthanasia. While acknowledging that it may be morally justifiable to administer a poison to some suffering patients who request it, they still oppose legalization of euthanasia because of the perceived social hazards.\(^4\)

Their apprehensions cover a wide range. They include fear of outright abuse. For example, some bioethicists worry that slanted presentations of the choices available will taint the patients' consent received, or that society will use euthanasia to eliminate socially isolated and unwanted individuals. Their apprehensions also include more subtle hazards supposedly flowing from the availability of active euthanasia: pressure on fatally stricken patients to accept death rather than undergo expensive life-sustaining therapy; diminution of medical efforts to palliate patients' suffering; weakening of society's commitment to care for the dying; erosion of professional medical mores; erosion of health-care providers' morale; and erosion of patient confidence in the medical profession flowing from concern about physician-caused death. Although these hazards may prove chimerical,\(^4\) Singer's book fails to address them.

One might respond to these various concerns by arguing that doctors rarely abuse their current role in withholding or withdrawing life support. Health-care providers have ample opportunity to exploit gravely afflicted patients by manipulating informed consent leading to life-support withdrawal decisions, by administering analgesics with an incidental effect of accelerating death and by removing life support from incompetent patients. In short, similar potential for abuse plagues both active euthanasia and withdrawal course. Of course, this may indicate only that suicide is more understandable and tolerable in some circumstances than in others.


of life support, and yet abuse has not materialized in the latter context.\textsuperscript{43}

I find this only partially reassuring. As I explain in Part III, the legal doctrine on the handling of incompetent dying patients and the standards for surrogate decisionmaking still are evolving. While medical professionals and surrogate decisionmakers have not perpetrated abuse on helpless patient populations, this fact may be attributable partially to the cautious evolution of legal standards for life-support removal. Those standards are still in flux and still need assessment.

Furthermore, advocates of active euthanasia cannot avoid the difficulties associated with surrogate decisionmaking simply by insisting that we confine active euthanasia to competent patients. Once we authorize active euthanasia, a natural impetus to extend its "benefits" to incompetent persons will follow. If a competent patient in unremitting pain would likely request and receive euthanasia, a strong impulse will emerge to extend the same benefit to an incompetent patient in a similar condition who never provided or never had the capability to provide advance instructions. Experience in the Netherlands confirms the existence of this impetus.\textsuperscript{44} Thus, proponents of active euthanasia ultimately must confront the issue of the standards for surrogate decisionmaking, a topic that Singer neglects except as to nonpersons.

III. \textbf{The Real Ethic: Autonomy and Constructive Preference}

Part I showed that Singer's supposed "old ethic," which mandates the preservation of human life except by extraordinary means, never prevailed. The real ethic of death and dying has developed in American jurisprudence over the past twenty years since the 1976 \textit{Quinlan} decision. Although some discontinuity between legal doctrine and medical practice persists, that jurisprudence has had a considerable impact on the professional standards applicable to end-of-life care.

Both the relevant cases and statutes of the past twenty years have tended to direct medical responses to fatal conditions according to patient preference, whether actual or putative. That autonomy-oriented thrust seems most evident when competent patients make contemporaneous decisions about medical intervention.

\textsuperscript{43} While many complaints are voiced about end-of-life medical practices, premature termination of life support is not one of them. \textit{See} sources cited \textsuperscript{infra} note 82.

\textsuperscript{44} Euthanasia has been performed there upon gravely incapacitated infants. \textit{See} \textsc{Carlos F. Gomez, Regulating Death: Euthanasia and the Case of the Netherlands,} 83-85 (1991); Maurice A.M. deWachter, \textit{Euthanasia in the Netherlands,} HASTINGS CTR. REP. Mar.-Apr. 1992, at 23, 24.
or when they have left advance instructions for their post-competence care. Yet, even when the now-incompetent patient never prepared such instructions, what I call constructive preference — a notion grounded in autonomy and respect for human dignity — emerges as the principal legal guide to the patient’s medical fate. The construction of preference requires a decisionmaker to project what the now-incompetent patient would want done. I call this approach constructive preference because it seeks to replicate a now-incompetent patient’s likely preference in the absence of actual patient choice. The surrogate’s decision inevitably must be constructive, but, as subsequent discussion will indicate, that decision need not be disconnected from the patient’s wishes.

An autonomy-constructive preference ethic does in fact underlie the current jurisprudence. Legal doctrine governing end-of-life medical care starts with the competent patient. American courts uniformly uphold the prerogative of competent patients to reject life-sustaining medical intervention.45 In so doing, they look to the doctrine of informed consent, a doctrine based on notions of bodily integrity and self-determination that rest, in turn, on respect for human dignity and capacity for choice.46 Thus, as to competent medical patients, a close relation exists between autonomy and dignity. The primacy of autonomy extends to “prospective autonomy” — a competent person’s right to shape her post-competence medical treatment by advance instructions. Numerous cases have looked to such instructions as the key determinant in surrogate decisionmaking.47

Legislatures also respect prospective autonomy. All fifty states accord statutory protection for some form of advance medical directive.48 Living will and advance directive laws, for example, give legal effect to the advance instructions of now-incompetent patients.49 Durable-power-of-attorney laws allow people to designate

46. See Beauchamp & Childress, supra note 1, at 142-46; Gerald Dworkin, The Theory and Practice of Autonomy (1988).
health-care agents to implement their advance instructions.\footnote{50} The apparent object of these legal sources — both judicial and legislative — is to respect self-determination in shaping one's own dying process.\footnote{51}

Some commentators challenge the notion of prospective autonomy.\footnote{52} They question the validity of choices made well in advance of incapacity and perhaps without a full understanding of and deliberation over the range of possible medical conditions and outcomes. They also contend that the interests that shape a declarant's advance instructions — in avoiding indignity, in avoiding the frustration of helplessness and debilitation, and in sparing loved ones from emotional and financial burdens — become largely irrelevant once incompetent patients no longer can appreciate violations of their prior choices.

Nonetheless, the overwhelming weight of judicial and legislative sentiment endorses prospective autonomy. The explanation is simple and understandable. People have a strong interest in shaping their own version of a dignified dying process regardless of whether they actually experience the feared degradation. Many adults have witnessed the demise of a loved one and can envision a level of debilitation that they deem intolerably undignified. People care about their lifetime image, which includes the memories left behind during the dying process.\footnote{53} They wish to imprint their values — whether grounded in religion, a personal vision of dignity, or solici- tude toward loved ones — on their end-of-life story. Prospective autonomy therefore protects important interests in self-definition and self-determination.\footnote{54}

Even if prospective autonomy is a meaningful concept, what about the incompetent person who never articulated her choices? Her surrogate's decision cannot invoke genuine autonomy — genu-

\footnote{50. \textit{Id.} at 1259-60.}

\footnote{51. "The principle of respect for persons, which supports respect for the autonomous patient's choices, also supports reliance on the nonautonomous person's prior autonomous directives." James F. Childress, \textit{Dying Patients: Who's in Control?}, 17 \textit{LAW MED. \\& HEALTH CARE} 227, 228 (1989).}


\footnote{53. \textit{See Cruzan v. Director, Missouri Dept. of Health}, 497 U.S. 261, 343-44, 356 (1990) (Stevens, J., dissenting); \textit{RONALD DWORmN, LIPE's DOMINION} 201-17 (1993); Harris, \textit{supra} note 23, at 6, 14.}

\footnote{54. For a more elaborate defense of prospective autonomy and discussion of its bounds, see \textit{NORMAN L. CANTOR, ADVANCE DIRECTIVES AND THE PURSUIT OF DEATH WITH DIGNITY} 23-32, 122-34 (1993).}
ine autonomy demands a personalized weighing of medical options in light of personal values and preferences. Nonetheless, the applicable legal norms in this context recognize a close relation between patient choice and surrogate decisionmaking. Both of the common standards governing decisionmaking on behalf of formerly competent patients — "substituted judgment" and "best interests" — display strong preoccupation with the patient's putative desires even when she never prepared advance instructions.

The substituted-judgment standard seeks to reach the same decision that the patient would reach if competent and cognizant of the circumstances. That objective — seeking to project and replicate what the patient would want — reflects the law's pervasive interest in respecting personal choice. Actual patient choice is impossible in the absence of prior instructions; however, the substituted-judgment approach serves the patient-choice goal by examining the patient's personal value system, including her "philosophical, religious and moral views, life goals, values about the purpose of life and the way [life] should be lived, and attitudes toward ... suffering and death." Some versions of substituted judgment authorize a surrogate decisionmaker to use such data in order to make a best approximation of what the patient would want, if competent.

The substituted-judgment standard thus seeks to treat the formerly competent patient as an individual with moral dignity whose putative preferences matter. By allowing a surrogate to consider a range of possible dispositions, from vigorous medical intervention to merely palliative care, it preserves the same range of options that would be available to a competent patient. In so doing, the substituted-judgment approach underlines the equivalence in stature between the now-incompetent person and her former competent self. Also, by striving to discern her likely wishes, the formula seeks to preserve the autonomy rights that the formerly competent patient no longer can exercise.

59. Notice how two of the earliest decisions involving incompetent patients stressed the goal of preserving a competent patient's right to choose: "The only practical way to prevent destruction of the right [to reject treatment] is to permit the guardian and family of Karen to render their best judgment... as to whether she would exercise it in these circumstances." In re Quinlan, 355 A.2d 647, 664, cert. denied, 429 U.S. 922 (1976). The Saikewicz court noted:
The problem with the substituted-judgment standard lies in its administration, especially in the absence of prior, considered instructions. Reliance on value-related or character-related data about the patient may engender uncertainty about what the patient would have wanted, if competent. To be sure, certain personal values may provide conclusive evidence of the patient's preferences. For example, a surrogate for a devout orthodox Jew or Roman Catholic who always has subscribed to her denomination's religious precepts safely can ascribe that denomination's well-developed positions regarding terminal care to the patient. Many factors commonly invoked to guide a surrogate under the substituted-judgment standard, however — such as the patient's prior attitude toward doctors, general lifestyle, and solicitude for the interests of close family — simply cannot identify the point of decline at which the patient would prefer death to continued existence.

Reliance on a patient's general value system, as part of a "best approximation" method for determining the patient's wishes, has prompted substantial criticism. The critics point to studies that indicate that a significant discrepancy exists between the wishes of seriously ill patients and the beliefs of their relatives as to what the patients would want. At best, these data suggest that some disjunction may lie between the surrogate's definition of the patient's wishes and the patient's actual, though unexpressed, wishes. At worst, they indicate that general value system or lifestyle indicia of patients' wishes leave room for surrogates to impose their own values and predispositions on patients. The ultimate specter is that surrogates, under the guise of the putative wishes of the patient, will make biased or self-interested determinations.

We think that principles of equality and respect for all individuals require the conclusion that a choice exists.... [W]e recognize a general right in all persons to refuse medical treatment in appropriate circumstances. The recognition of that right must extend to the case of an incompetent, as well as a competent, patient because the value of human dignity extends to both.


62. See Dresser & Robertson, supra note 52, at 235; Lo, supra note 60, at 216; Tracy L. Merritt, Equality for the Elderly Incompetent, 39 STAN. L. REV. 689, 709, 714 (1987); Rhoden, supra note 60, at 387.
The solution, however, is not to abandon the substituted-judgment standard's focus on putative patient wishes. Rather, it is to guard against excessive surrogate subjectivity by formulating default norms to guide and circumscribe surrogate decisionmaking.63

Beside substituted judgment, the other major standard for surrogate decisionmakers is the “best interests of the patient.”64 Under one version of best interests, the surrogate seeks to assess the objective well-being of the now-incompetent patient and maintains life support unless the prospective burdens on the patient — primarily pain and suffering — appear to outweigh the benefits — pleasure and satisfaction.65 A best-interests formula, however, can and usually does encompass more than just the observable emotions of the patient.

In its own fashion, a best-interests standard impels the surrogate to effectuate what the now-incompetent patient would have wanted, if competent. In the absence of proof about the patient’s actual wishes, the best-interests standard assumes that the patient would want the same treatment that the average person in the same circumstances would want. It defines patient well-being — the key to best interests — according to understandings about the average person’s definition of well-being.66 Extreme suffering, for example, is regarded as an integral ingredient of best interests because the vast majority of people are averse to extreme suffering. Quality of life, often addressed under the rubric of patient dignity, frequently forms an element of best interests because the average person ties the two together.67 The hope is to implement the patient’s likely

63. For elaboration on this idea, see infra notes 72-78 and accompanying text.

64. Many view the two standards — substituted judgment and best interests — as part of a continuum. The surrogate starts with substituted judgment and seeks to ascertain what the patient would have wanted by considering prior expressions and other indicia. When those indicia prove indeterminative, the surrogate attempts to define the best interests of the patient. See Phillip G. Peters, Jr., The State’s Interest in the Preservation of Life: From Quinlan to Cruzan, 50 OHIO ST. L.J. 891, 922-23 (1989); Stewart G. Pollock, Life and Death Decisions: Who Makes Them and By What Standards, 41 RUTGERS L. REV. 505, 518-22 (1989); Robert M. Veatch, Forgoing Life-Sustaining Treatment: Limits to the Consensus, 3 KENNEDY INST. OF ETHICS J. 1 (1993).

65. See Dresser, supra note 52, at 657 n.2, 711; Rhoden, supra note 60, at 398-99.


67. See NEW YORK STATE TASK FORCE, supra note 41, at 77-113 (arguing that life support seems excessively burdensome for a “patient who would have viewed continued treatment as an affront to his or her dignity”); see also BEAUCHAMP & CHILDRESS, supra note 1, at 219. Clinical protocols — guidelines for end-of-life care prepared by professional organizations — often include quality of life as an element of best interests. See AM. MEDICAL ASSN., CURRENT OPINIONS OF THE COUNCIL ON ETHICAL AND JUDICIAL AFFAIRS, § 2.16;
choice by having the surrogate use the same criteria and weighting of factors that most people would choose for themselves. The best-interests and substituted-judgment approaches thus have a common objective — replication of what the individual patient likely would want regarding end-of-life care.

While the best-interests and substituted-judgment formulae share a common ethic, they also share common difficulties. Both raise the same kinds of concerns — indeterminacy and potential subjectivity in surrogate decisionmaking. For example, to the extent that the best-interests standard considers suffering as a key element, discerning the experiential reality of gravely demented patients seems a daunting, if not impossible, task. Indeed, severe problems of measurement plague any surrogate seeking to determine an incompetent's level of suffering or to compare her levels of suffering and satisfaction. Likewise, imprecision nags at any quality-of-life determination as an ingredient of best interests. Some commentators dismiss quality of life — or indignity — as a subjective, value-laden notion that lacks consistency and falls prey to surrogates' biases regarding a minimally tolerable quality of life.

In sum, the jurisprudence of surrogate decisionmaking strives to implement the actual or putative wishes of incompetent patients. Yet if we really want to implement those wishes, we must overcome


68. See The Hastings Center, Guidelines on the Termination of Life-Sustaining Treatment and the Care of the Dying 28 (1987).

69. The two standards do not ignore the actual preferences of the particular patient. The substituted-judgment standard remarks them before turning to a best-interests approach, and the best-interests formula often consults the patient's discernible values and preferences in defining her best interests.

70. "The real burdens and benefits of life in extremely debilitating circumstances are often beyond our ability to know confidently or comprehend fully." Peters, supra note 64, at 942. For a comprehensive examination of this issue, and one urging greater efforts to discern the murky reality in question, see Dresser, supra note 52, at 666-91. Beside Professor Dresser, many other commentators have noted the intrinsic difficulty in measuring the burdens and benefits of a severely demented patient. See John Arras, The Severely Demented, Minimally Functional Patient: An Ethical Analysis, 36 J. Am. Geriatric Soc. 938 (1988); Rhoden, supra note 60, at 404-05; Nancy K. Rhoden, The Limits of Legal Objectivity, 68 N.C. L. Rev. 845, 847-50 (1990); see also In re Peter, 529 A.2d 419, 424-25 (N.J. 1987).

the imprecision and subjectivity inherent in concepts such as intolerable indignity or quality of life. This is essential because any effective surrogate decisionmaking standard must reassure competent persons that their post-competence care will conform to their desires and expectations. Unless we achieve a common understanding of intolerable indignity, for example, a decisionmaking standard that incorporates that element will engender anxiety rather than reassurance. We need, therefore, reliable guidelines about levels of intolerable indignity to serve as a check on arbitrariness and abuse in surrogate decisionmaking.

I use the term "constructive preference" to denote an approach that surrogates may employ when making end-of-life medical decisions for formerly competent patients who left no instructions. The object — as with much of substituted-judgment and best-interests doctrine — is to provide the medical care that the now-incompetent patient would have chosen if she had considered the issue while competent. Because the patient never exercised her prospective autonomy prerogative or provided definitive guidance with regard to her end-of-life treatment, the surrogate will do her best to determine what most people would want in the same circumstances and to treat the patient accordingly. As Nancy Rhoden argues, acting on a patient’s "probable desires can be equated with implementing the patient’s right of choice.”

Constructive preference rests on the premise that most people want to avoid extreme indignity in their own post-competence dying processes. Constructive preference also assumes that widespread accord exists about intolerable levels of debilitation in the dying process, and that this accord will allow for some default

---

72. Sanford H. Kadish comments:

How much ability to sense and take comfort from experiences is required before we can say [a debilitated patient’s] life is not worth living? At bottom, the difficulty is that we have no way to make confident judgments about how far cognitive and physical deterioration must go before life ceases to be worth living, because the value judgments implicit in such a conclusion are in sharp contention in our society.


Rebecca S. Dresser also cites the "highly disparate meanings" dignity can have for different people. Rebecca S. Dresser, Life, Death, and Incompetent Patients: Conceptual Infirmities and Hidden Values in the Law, 28 ARIZ. L. REV. 373, 387 (1986).


74. Under the prevailing autonomy ethic, a patient's actual preferences, when discernible, should govern. Therefore, constructive preference provides a fallback when a patient's history and values provide no definitive guidance to her surrogate.

75. Rhoden, supra note 60, at 384.
guides. Of course, the array of circumstances that confront dying patients is enormous and not every patient’s situation can be resolved by resort to widespread accord or consensus. At least for some commonly confronted circumstances, though, surrogates may find guidance in people’s widely shared predilections about intolerable levels of debilitation.

Permanent unconsciousness provides the best example of consensus sufficient to trigger constructive preference. Surveys consistently show that the vast majority of people would not wish to have life support to maintain them in a permanently unconscious state. Given this, a surrogate should be required to authorize the cessation of life support for a permanently unconscious patient absent significant evidence that the patient’s views deviate from the common preference.

The constructive-preference approach raises many issues: How can we measure common preferences about indignity, given the multitude of potential death-and-dying circumstances? Whose preferences should matter in establishing a norm? What impact upon a surrogate’s choice should flow from the fact that x or y percentage of people deem a particular status intolerably undignified for their own future fates?

I address only the first question here. As to data sources, people’s preferences regarding post-competence medical care can be gleaned from surveys and from bulk analysis of advance medical directives. While some advance directives seem cursory and uninformative, others spell out clear visions of intolerable indignity in the dying process. Also, although surveys cannot anticipate the

76. Several commentators recognize the need for default positions, grounded on understandings about what most people would want for themselves, to guide decisions on behalf of incompetent patients who have not left sufficient indicia of their personal preferences. See James F. Drane & John L. Coulehan, The Best-Interest Standard: Surrogate Decision Making and Quality of Life, 6 J. CLINICAL ETHICS 20, 24-26, 29 (1995); James Lindgren, Death by Default, LAW & CONTEMP. PROBS. 1993, at 185, 186, 195, 199, 228-29; Carl E. Schneider, From Consumer Choice to Consumer Welfare, HASTINGS CTR. REP. Nov.-Dec. 1995, Special Supp., at S25, S27 (urging default positions for patients based on “what we think they would want if they thought about it”).

77. See, e.g., Lindgren, supra note 76, at 231.


79. See Emanuel & Emanuel, supra note 61, at 6; Lindgren, supra note 76.

80. See Norman L. Cantor, My Annotated Living Will, 18 LAW MED. & HEALTH CARE 114 (1990). Some advance directives utilize values histories or values profiles to provide guidance about intolerable levels of debilitation. See CANTOR, supra note 54, at 166-70; Ezekiel Emanuel & Linda L. Emanuel, Living Wills: Past, Present, and Future, 6 J. CLIN. ETHICS 9, 15-16 (1990); Linda L. Emanuel, Structured Deliberation to Improve Decisionmaking for the Seriously Ill, HASTINGS CTR. REP. Nov.-Dec. 1995, Special Supp., at S14; Pam
multitude of circumstances that will confront incompetent patients, they can utilize scenarios that reflect a range of commonly occurring conditions in the dying process.\textsuperscript{81}

Constructive preference assumes, in the end, that default presumptions can be anchored in objectively measurable data about the level of mental and physical debilitation that most people consider intolerably undignified and therefore unacceptable. It assumes that we can establish guidelines or presumptions for certain commonly occurring conditions. When a large majority of people would prefer withdrawal of life support, a surrogate should implement the popular preference and withdraw life support, unless significant indicia in the particular patient’s history indicate that the patient would prefer otherwise. Default principles would have to receive wide publicity, so that any person whose preferences differed from the default position could issue advance instructions and avoid imposition of constructive preference. By focusing on what competent people commonly choose and reject, constructive preference discourages resort to surrogates’ subjective visions about which lives are worth preserving or to any government-formulated view of minimally acceptable dignity. Moreover, by following a course that the majority of people would choose — that is, implementing the course that the now-incompetent patient would likely have chosen — constructive preference comes as close as possible to fulfilling the wishes of people who have never communicated their wishes or left other meaningful indicia of their preferences for end-of-life medical intervention.

\textbf{Conclusion}

Contrary to what Singer suggests, an ethic already exists in the context of death-and-dying that does not adopt a maximum extension-of-life principle. The ethic that permeates the past twenty years of jurisprudence places primacy on autonomy — both contemporaneous and prospective — and on constructive preference when a patient’s actual preference cannot be determined. The contemporary ethic recognizes quality-of-life distinctions that are grounded in competent persons’ choices regarding intolerable in-

\textsuperscript{81} See, e.g., Linda L. Emanuel et al., \textit{Advance Directives for Medical Care — A Case for Greater Use}, 324 New Eng. J. Med. 889 (1991); Linda L. Emanuel & Ezekiel Emanuel, \textit{The Medical Directive}, 261 JAMA 3288 (1989); Emanuel & Emanuel, supra note 61, at 6. Both scenarios and values profiles — documents asking people to identify elements of personally intolerable indignity — will permit us to learn about common attitudes toward end-of-life care.
dignity in the dying process or in surrogates’ choices based on understandings of what most people would wish in similar circumstances. The challenge is not, as Singer claims, to propound a radically new ethic — although he may be right that active euthanasia ultimately will be added as an available option. Rather, the challenge is to translate the extant theory into practice. For in American institutions the sad reality continues to be that the dying process often is not what patients want — or would have wanted.\footnote{82. A large recent study seems to indicate that the dying process in many hospitals still is characterized by absence of communication between patients or surrogates and caregivers, misunderstanding about the wishes of the patient, and over-commitment to aggressive intervention. See The Study to Understand Prognosis and Preferences for Outcomes and Risks of Treatment (SUPPORT), \textit{A Controlled Trial to Improve Care for Seriously Ill Hospitalized Patients}, 274 JAMA 1591 (1995). For reactions to this study and suggestions about how to conform customary practice to ethical theory, see Bernard Lo, \textit{Improving Care Near the End of Life: Why Is It So Hard?}, 274 JAMA 1634 (1995); \textit{Dying Well in the Hospital: The Lessons of SUPPORT}, \textit{HASTINGS CTR. REP. Nov.-Dec. 1995}, Special Supp., at S1-S36.}
SEARCHING FOR POSITIVISM

Philip Soper*


INTRODUCTION

Issues in philosophy have life cycles of their own. Problems neglected or forgotten revive — either in response to the times (witness the renewed interest in political theory during the Vietnam War) or in response to a new and original argument (the revival of Kantian ethics after Rawls). Once revived, these issues often enjoy a brief period of intense flourishing, followed by a gradual return to dormancy — either because the times move on to make the issue less pressing, or because a point of diminishing originality or interest is reached in academic discussions.

The nature-of-law issue in jurisprudence might seem to be in the late stages of one of these periodic cycles. Revived more than thirty years ago with the appearance of H.L.A. Hart's The Concept of Law¹ and energized by Ronald Dworkin's original reformulation of the "natural law" side of the issue,² the entire field flourished over the next quarter century in a way that was both unprecedented and bound to lead to subtle and enlightening refinements.

In view of the extensive interest in this issue in recent years, one understandably might think that we are now close to exhaustion and that little more can be expected until the next turn of the cycle. All the more pleasant to discover Professor Waluchow's³ fine book. Waluchow's contribution is original and beautifully crafted, and it also provides one of the best overviews of the debate during the past thirty years. The writing style is enviable for its clarity, and the argument is admirably honest, giving fair due to opposing views and anticipating and handling virtually all serious objections to the argument that might occur to any reader. Although the book is

---


3. Associate Professor and Chair, Department of Philosophy, McMaster University.
pitched at a fairly high level — one that seems to take for granted reader familiarity with much of the literature in the area — the result for those who follow this field is a rich and rewarding tour of the landscape. Despite the familiar themes and issues, Waluchow never fails to place his own stamp on them by the way he summarizes, clarifies, or refines the ideas. The book deserves a place on anyone's jurisprudence shelf.

What makes Waluchow's book so intriguing, in part, is that it purports to aim at a very narrow question — a small gap, one might think, in positivist legal theory. Waluchow indeed is refreshingly, if unnecessarily, modest in confessing to a "fairly narrow scope" and a "lack of novelty" in his thesis (p. 4). But the gap he aims to fill, as is often the case, proves to be a chink through which one sees, with the aid of Waluchow's analysis, deep inside the natural law-positivism debate. As a result, one comes away with a clearer view of the entire field as well as a powerful argument for a new and arguably better form of positivism.

The small gap in the positivist literature results from an interesting twist on the usual question that underlies these debates. The question usually posed is whether there is a necessary (conceptual) connection between law and morality; Waluchow asks instead whether there is a necessary lack of connection between these concepts. This shift in focus reflects developments in legal theory since Hart's book first appeared. Whereas Hart suggested that connections between law and morality were, at best, contingent rather than necessary, subsequent developments in the positivist literature, supported by characterizations of positivism from the natural law side, support a stronger thesis: Far from there being no necessary connection between law and morality, there is a necessary separation of law and morality; moral standards cannot be part of the "law" in legal systems.

Waluchow calls this strong thesis, most clearly associated with the work of Joseph Raz,⁴ "exclusive positivism." He sets out, in contrast to Raz, to defend "inclusive positivism" — the view, originally suggested by Hart, that law may include moral standards, but need not do so. In the course of the argument, Waluchow also presents two other theses: (1) Inclusive positivism is distinguishable from and superior to Dworkin's alternative, nonpositivist account of how moral principles figure in law; and (2) Although it is a contingent question whether a legal system includes moral principles among the legal standards courts are asked to apply, it is in general a good thing for them to do so.

Along the way to establishing these claims, Waluchow develops a number of original distinctions and insights into the current debate — too many to be summarized usefully here. In this review, I focus on the central argument of the book — whether and why positivism can or cannot accept moral principles as part of the law and, assuming it can, whether the case for positivism over natural law is thus strengthened. In brief, I argue: (1) that Waluchow does establish a case for viewing inclusive positivism as superior in some respects to exclusive positivism; but (2) that this victory is only partial — neither version of positivism, exclusive or inclusive, is quite good enough. The central idea of the natural law theorist concerning the conceptual connection between law and morality survives the challenge of both forms of positivism because that connection is of a different — and more modest — kind than commonly has been assumed.

I. THE BACKGROUND

The question about the status of moral principles in legal theory began as a by-product of the debate that followed Dworkin's challenge to Hart.5 Dworkin stressed that the “right answers” to legal decisions were, in theory, to be found by undertaking a Herculean inquiry in two dimensions: (1) institutional history — convention and legal norms identified by the ordinary positivist's pedigree (the "fit" dimension); and (2) the political and moral theories that provide the best justification for using those conventions as the basis for state coercion (the "moral" dimension). Dworkin's theory encountered two initial responses designed to suggest that the theory might, after all, be just another form of positivism. First, because the requirement of "fit" seemed to restrict Dworkin's moral principles to the conventions of the particular society in which judges found themselves (conventional morality, not true morality), some claimed that Dworkin had advanced only a more refined version of positivism: conventional moral principles should be added to the positivist's conventional rules as part of the "pedigree" or test for determining legal norms.6 Second, even if Dworkin was right that the moral dimension required a judge to reach beyond conventional morality into true principles of political morality, that fact was itself the result of social conventions authorizing or requiring judges to adopt this particular method of adjudication. Thus, just as Hart had suggested that legal systems contingently could include moral standards (as in the Due Process Clause of the United States Constitu-

5. See Dworkin, Hard Cases, supra note 2; Dworkin, The Model of Rules, supra note 2; Dworkin, Social Rules and Legal Theory, supra note 2.

tion), so Dworkin's theory could be viewed as reflecting an assignment to judges in Anglo-American jurisdictions to test laws by reference to political morality — a contingent fact about Anglo-American jurisprudence rather than a necessary feature of all legal systems.\(^7\)

It was Dworkin's response to this second suggestion for putting a positivist gloss on his theory that led to the first explicit statement of the thesis that Waluchow calls "exclusive positivism." Positivism, according to Dworkin, could not invite judges simply to use moral standards to decide cases and still remain positivism. That is because positivism

is connected to a more general theory of law — in particular to a picture of law's function. This is the theory that law provides a settled, public and dependable set of standards for private and official conduct, standards whose force cannot be called into question by some individual official's perception of policy or morality.\(^8\)

Two features about morality, according to Dworkin, prevent a positivist from counting moral standards as part of the law: first, because moral standards are inherently controversial, they could count as "law" only by abandoning the positivist's central thesis about law's essential function; second, even if one thinks moral standards rest on objectively determinable "moral facts," that claim about the status of moral judgments is itself controversial; the whole point of positivism, according to Dworkin, is to provide a theory of law that is "independent of any controversial theory either of metaethics or of moral ontology."\(^9\)

As noted, Dworkin's claim about the legal status of moral standards under positivism initially was tendered simply as a response to critics of his theory; it was not the result of a full-fledged argument about the "essence" of positivism. But, then, neither was Hart's contrary suggestion. Hart's remark that legal norms could incorporate moral standards seemed almost an offhand, casual way of responding indirectly to some of the arguments of his main antagonist at the time, Lon Fuller. Fuller's emphasis on the role of purpose and reason in the interpretation and application of legal rules would be consistent with the denial of a necessary connection between law and morality if the appeal to such standards were sim-
ply the result of a contingent incorporation of morality in legal provisions.10

In his later work, Dworkin repeated his claim about the essence of positivism. By now the version of positivism that Dworkin thought inconsistent with the “essence” of positivism had assumed a new name — “soft conventionalism.” “[S]oft conventionalism instructs judges to decide according to their own interpretation of the concrete requirements of legislation and precedent, even though this may be controversial . . . . [S]oft conventionalism is not really a form of conventionalism at all . . . . It is, rather, a very abstract, underdeveloped form of law as integrity.”11 Despite the new name, Dworkin’s defense of the view that soft conventionalism could not be a “true” form of positivism remained, at that point in the debate, still too cursory to be much more than an assertion rather than a full-fledged argument.12 Moreover, although Hart never returned publicly to the debate before his death, the postscript published with the most recent edition of The Concept of Law seems to confirm, in some respects, Hart’s continued endorsement of a more relaxed form of positivism — a form that Hart called “soft positivism.”13 On the question whether standards could count as “law” if they were too controversial or uncertain, Hart seemed to say that the question was one of degree: legal systems could tolerate some uncertainty and still provide the kind of guidance and social control that was an important, though not necessarily a “paramount and overriding,” concern for a positivist theory.14 As for Dworkin’s second objection — that positivism required a theory of law that remained independent of controversial questions about the status of moral judgments — Hart seemed to agree: “I still think legal theory should avoid commitment to controversial philosophical theories of the general status of moral judgments and should leave open, as I do in this book, the general question of whether they have what Dworkin calls ‘objective standing.’ ”15

10. Compare HART, supra note 1, at 204 (making the point about legal systems incorporating moral principles: “If this is what is meant by the necessary connection of law and morals, its existence should be conceded”) with id. at 207 (referring explicitly to Fuller’s “inner morality of law”: “Again, if this is what the necessary connection of law and morality means, we may accept it.”).

11. DWORm, LAW’s EMPIRE, supra note 2, at 125, 127-28.

12. For a critique of Dworkin’s refusal to recognize soft conventionalism as a form of positivism, see Philip Soper, Dworkin’s Domain, 100 HARV. L. REV. 1166, 1177-79 (1987) (reviewing DWORm, LAW’s EMPIRE, supra note 2).

13. See HART, supra note 1, at 250-54. It is not clear whether this second edition of Hart’s book, which was published in 1994 with the additional postscript, was available to Wuluchow before the appearance of his book.

14. See id. at 252.

15. Id. at 253-54 (citation omitted).
It remained to Joseph Raz to provide, in contrast to this rather casual exchange of views between Dworkin and Hart, the first fully sustained argument in defense of the view that law could not include moral principles among the tests for legal validity. Raz called this thesis the "strong social thesis" or, simply, the "sources thesis." As Waluchow characterizes it, this is the view that the existence of a valid legal rule is solely a function of whether it has the appropriate source in legislation, judicial decision or social custom, matters of pure social fact, of pedigree, which can be established independently of moral factors. In addition, the content of a legal rule can be determined, Raz believes, by establishing facts about human beings (e.g. their legislative actions and intentions) that can be ascertained without the use of moral arguments. [p. 82; footnote omitted] In defending this view, Raz affirmed Dworkin's vision about the essence of positivism but, of course, argued that his own vision was a better theory of law than its alternatives — particularly Dworkin's. The point of rehearsing this background is that it enables one better to appreciate Waluchow's contribution. Like Raz, Waluchow also provides a full-fledged inquiry — the only one apart from Raz's of which I am aware — into the question whether law can include moral standards. Unlike Raz, Waluchow endorses Hart's original suggestion: soft positivism — "inclusive" positivism for Waluchow — is a viable form of positivism and preferable to Raz's exclusive form. But whereas Raz's endorsement of exclusive positivism meant that he did not need to confront Dworkin's claim that soft positivism was really not positivism at all, Waluchow cannot avoid Dworkin's challenge. Thus, he devotes a significant portion of his book to explaining why, contrary to Dworkin's view, inclusive positivism does not collapse into a form of natural law or law-as-integrity. It is to these two arguments that I now turn.

II. EXCLUSIVE POSITIVISM

A. The Central Argument

John Austin began his famous lectures on jurisprudence with the observation that law, in its most general sense, "may be said to be a rule laid down for the guidance of an intelligent being by an intelligent being having power over him." Though he hardly could have

16. See Raz, supra note 4, at 47.
17. For an earlier, but briefer, exploration of this question, see Philip Soper, A Theory of Law 101-09 (1984).
18. Raz's view on this issue was that soft positivism "is on the borderline of positivism and may or may not be thought consistent with it." Raz, supra note 4, at 47.
anticipated the debate that was to occur 150 years later over the implications of his theory, Austin's simple statement indirectly explains the appeal of exclusive positivism. To count as law, a standard at a minimum must be able to serve as a guide to human conduct. The point has less to do with the peculiarly "legal" than it does with the prerequisites for speaking of a system of social control in the first place. Only standards that are reasonably able to convey to subjects what they are expected to do can serve as means of social control.

Although Waluchow confronts this basic claim about the function of law head on, he does not always seem to appreciate its force. In part that is because he is, if anything, too catholic in his willingness to examine each and every argument for exclusive positivism made by, or implied in the writings of, its main proponent, Joseph Raz. Thus Waluchow confronts the "linguistic argument," the "argument from bias," the "institutional connection argument," the "argument from explanatory power," the "argument from function," and the "authority argument" (pp. 103-40). While Waluchow does not suggest that each of these arguments is equally plausible — and, by and large, his critique of these arguments is persuasive — the effect of his wide-ranging survey and meticulous criticism is to detract from the central argument, the argument from function.

To illustrate how the central argument eclipses the others, consider, for example, the argument from authority, to which Waluchow devotes the most attention. Raz believes that law claims authority, even though it does not always have authority. In order even to claim authority, however, legal systems must satisfy the logical preconditions for authority. Relying on an analogy to arbitration, Raz explains that these conditions entail that when the authority makes a decision, the factors that went into that decision are then excluded from reconsideration by those who are subject to the authority. The authority represents the "executive" stage of decisionmaking that follows deliberation. There could be no authority if those subject to the authority could go behind the decision to reconsider the reasons on which the decision was based. So, too, with law and morality. Whenever an institution, judge, or legislator makes a decision based on moral factors, those underlying moral factors are excluded from the factors that "guide" the citizen; what remains as "law" is the decision itself — the legal norm or rule identified by the social fact of the decision.

Waluchow's discussion of this argument (pp. 129-40), which examines in some detail Raz's account of exclusionary reasons as it relates to his concept of authority, seems an unnecessary diversion for two reasons. First, it is not clear that law does claim authority in

Raz's sense.\textsuperscript{21} If law makes no such claim, or in any event if this issue is controversial, then the argument for exclusive positivism — to the extent that it depends on insisting that law claims authority — will be equally controversial. Second, and more to the point, it does not matter whether or not law claims authority because before law can do so it first must explain what is to be done. It first must provide guidance about the act to be done before it can attach to that act any further claims about the spirit in which the act is to be done, for example, in response to a claimed obligation to obey, or simply in order to avoid a sanction. Guiding action, in short, must come first; if standards are incapable of doing this, they seem a fortiori incapable of serving as the basis for a claim of authority, or a command, or a norm, or a piece of advice.\textsuperscript{22}

The metaphor that has figured in so much of the debate about the nature of law — that of the gunman — illustrates the point. Imagine a mugger who gives me the following order: "Hand over your money if it is the just thing to do." Should we say that this gunman has given a rather vague order? Or should we say instead, as the exclusive positivist presumably would, that this gunman has not yet given any order at all? If one thinks that the simple instruction "do justice" is too indefinite to provide a guide to expected conduct, several remedies are possible. The most obvious is to await further clarification by the speaker. Once a decision is made — in the above case by the gunman about what \textit{he} thinks I should do, or what \textit{he} thinks justice requires — doubt about what is expected dissolves. Another possibility, which requires switching from the metaphor of the gunman to an analogy with common law courts, is this: repeated decisions about what "justice" requires may begin to accumulate in a way that permits extrapolating from past cases to a conclusion about what consistency would require in the instant case.\textsuperscript{23} In either case one needs something more — clarification or context — before one can tell what is expected.

\textsuperscript{21} For a defense of the view that law makes a weaker normative claim than that entailed by the "claim of authority," see Philip Soper, \textit{Law's Normative Claims, in THE AUTONOMY OF LAW} (Robert George ed., forthcoming 1996).

\textsuperscript{22} As another recent supporter of Raz's view suggests, "law is the determinations of authorities of what ought to be done." Larry Alexander, \textit{All or Nothing at All? The Intentions of Authorities and the Authority of Intentions, in LAW AND INTERPRETATION: ESSAYS IN LEGAL PHILOSOPHY} 357, 359 (Andrei Marmor ed., 1995).

\textsuperscript{23} In previous work, I suggested that one could view the common law in just this way. Starting from the simple rule of recognition that "[a]ll disputes are to be settled as justice requires," one might build up a body of case law that, together with the requirement of consistency (treat like cases alike), would produce a fairly determinate system of law. \textit{See} Soper, \textit{supra} note 7, at 512 & n.129. For an arguably similar view of the common law, see Stephen R. Perry, \textit{Second-Order Reasons, Uncertainty and Legal Theory}, 62 S. CAL. L. REV. 913, 953-93 (1989). Waluchow here follows Dworkin in rejecting such a rule of recognition as too indeterminate to function as a "legal system." \textit{See} p. 185. In doing so, he ignores the point that the system could become determinate over time. He also seems to ignore his own willingness elsewhere to accept that moral standards may be more determinate depending on
B. Critique of the Central Argument

Waluchow considers the argument from function as it appears both in Dworkin's claims about the essence of positivism as well as in Raz's more elaborate theory. In both cases, his response is essentially the same. First, Waluchow claims that even if relative certainty in guiding conduct is an important goal of law, it need not be viewed as the "essential" or only goal; law can serve other functions. For example, law may have "educative" functions (p. 132) that accept some uncertainty in return for the value of having judges and litigants think and argue in substantive moral terms (pp. 121-22, 134-35). Second, even if relative certainty is a major characteristic of legal standards, some moral standards are sometimes relatively certain, or at least no more uncertain than some of the social facts that the positivist counts as law.

The first of these objections leads to the problem of definition that has always haunted this field: How are we to resolve disputes about the "essence" of a concept and, thus, about whether the concept of law is "essentially" connected to the concept of certainty?\textsuperscript{24} I shall return later to this objection. For now, it is the second objection that most directly challenges the sharp distinction that the exclusive positivist draws between moral and legal standards. That distinction assumes that a clear conceptual or practical difference regarding certainty justifies distinguishing social facts from moral standards and warrants allowing only the former to serve as candidates for legal standards.\textsuperscript{25} In order to assess this issue, it may help to confront an ambiguity that hinders the discussion: What kind of uncertainty is at stake?

1. Uncertainty and Indeterminacy

Waluchow notes that there is a distinction "between uncertainty (an epistemic property) and indeterminacy (a logical property)” (p. 238). Standards that are indeterminate do not admit of “cor-
rect" answers, even in theory; thus indeterminacy entails uncertainty. But the converse is not true; a standard whose application is uncertain may, nevertheless, have a single correct answer, even though that answer is controversial and difficult to ascertain.

Which of these concepts is the critical one for theories of law that stress the importance of law's ability to guide conduct? If legal standards must be able to answer questions about what one is expected to do, one might think that uncertainty is the critical concept. Standards that fail to guide, either because they admit of no single correct application or because the correct application is inherently controversial, are equally incapable of constraining judges and guiding citizens. In both cases, then, a judge applying such standards must be making law, not finding it.

Although he is not always clear on this point, Waluchow seems to accept indeterminacy rather than uncertainty as the critical, limiting concept in deciding whether we are dealing with "law." The relationship between these two concepts and Waluchow's views on its implications for a positivist theory of law can be found in the following three conclusions that he seems to support.

First, indeterminacy is inevitable in any legal system, both because some moral standards that one tries to enact as law may have no determinative answer (pp. 186, 223) and because of the open texture of language in most legal standards. In these cases, there is no law until a judicial decision is made, and judges in such cases must be seen as legislating.

Second, moral standards do "sometimes provide correct and uncontroversial answers to questions of legal validity" (p. 226). In such cases, moral standards can be counted as law, consistent with the function of providing guidance.

Finally, as to moral questions that might be thought to be inherently controversial, these standards could also qualify as law for a positivist as long as they still can be said to have a right answer (p. 224). In other words, only those moral standards that entail indeterminacy fail to qualify as law; uncertainty alone is not disqualifying. It is this last conclusion that seems to indicate that Waluchow — like Dworkin, on whose arguments Waluchow here relies — views indeterminacy, rather than uncertainty, as the critical factor governing whether moral standards can count as law. As to the objection that uncertain standards, even if they have right answers, cannot provide the guiding function critical to law, Waluchow's re-

26. Waluchow, in a sensible discussion of Hart's views on the issue, suggests that the question of whether standards are determinate is a matter not only of the "plain meaning" of the language, but of purposes and background understandings that often supplement language to yield determinable results. In the end, however, even with these supplemental, interpretive aids, "indeterminacy will be encountered somewhere along the line." P. 250.
sponse seems to be that guidance, though important, is not that im-
portant — it is not essential to the concept of law.

C. Positivism and Metaethics

Whether readers will agree with Waluchow's claim that moral
standards can be certain enough to count as law will, of course, de-
pend on one's views about the underlying problem of moral philos-
ophy: What is the status of moral judgments? That question, in
turn, leads to a more basic puzzle in connection with Waluchow's
attempt to defend moral standards as legal standards in a positivist
account: What has happened to the second objection, raised by
Dworkin, according to which positivism must be seen as committed
to a theory of law that is independent of the controversial issues of
moral ontology and metaethics?27

It now appears that we have three different claims in the litera-
ture about the relation between positivism and metaethics. The
first is a claim of complete independence of the sort originally made
by Dworkin and developed by Raz — that positivism must remain
"independent of any controversial theory either of meta-ethics or of
moral ontology."28

The second position appears to be the position Hart endorses:

I still think legal theory should avoid commitment to controversial
philosophical theories of the general status of moral judgments and
should leave open ... the general question of whether they have what
Dworkin calls 'objective standing.' ... Of course, if the question of
the objective standing of moral judgments is left open by legal theory,
as I claim it should be, then soft positivism cannot be simply charac-
terized as the theory that moral principles or values may be among
the criteria of legal validity, since if it is an open question whether
moral principles and values have objective standing, it must also be an
open question whether 'soft positivist' provisions purporting to in-
clude conformity with them among the tests for existing law can have
that effect or instead, can only constitute directions to courts to make
law in accordance with morality.29

Hart's position here is somewhat curious: it leaves the question
of the status of moral principles open, but then it seems to accept as
a consequence that one must also leave open the question whether
these principles really are law. Far from being independent of
metaethical issues, this view seems to lead to a legal theory that is

27. Although "metaethics" is often used to refer to inquiries into the meaning of moral
terms, I shall use it to refer generally to the entire related range of problems in moral theory
that Dworkin originally suggested a positivist theory must avoid: controversial questions
about moral ontology (the status of moral judgments) as well as questions about the meaning
and application of moral terms.

28. DWORKIN, TAKING RIGHTS SERIOUSLY, supra note 7, at 349.

29. HART, supra note 1, at 253-54 (citation omitted).
dependent on the resolution of metaethical issues; the theory remains aloof from the question of the status of moral judgments, but as a result the legal theory itself will remain incomplete until that issue is settled.

In contrast to both of these positions, Waluchow seems to claim that positivism need be neither independent of, nor aloof from, questions of metaethics. Instead, as noted, Waluchow simply asserts that moral standards are sometimes determinate and capable of guiding conduct, thus apparently committing himself to a position in favor of the objectivity of moral judgments.

The problem with Waluchow's position is that it seems to require arguments that the book really does not contain, arguments defending the objective status of moral judgments. Unfortunately, the most that we get in the way of argument is the assertion, at various points, that "not all moral questions are inherently unsettled" (p. 115). As evidence for this assertion, Waluchow cites the fact that no one could doubt, for example, that a statute purporting to enslave citizens would be held to violate constitutional or Charter provisions that protect equality and similar "moral" rights (p. 115). But this evidence does not help Waluchow. When there is a clear consensus on a moral issue, the exclusive positivist can also accept the Constitution's reference to equality as a legal standard because it refers to a social fact, a fact about people's beliefs.

It seems that Waluchow is in something of a quandary. If he relies on well-settled moral views to "prove" that moral standards can serve as legal standards, he only will aid the exclusive positivist by appearing to make social facts the test for law. If, on the other hand, he claims that moral standards can count as law even when they are not well-settled, he must defend the apparent willingness to claim objectivity for such standards despite their controversial nature — a task in metaethics that he does not undertake.

There is, I think, a possible way out of this dilemma, but it is one that requires a return to a position of independence or aloofness from metaethical questions. I have been suggesting that Waluchow takes a position on the controversial question of the objectivity of moral judgments. In fact, however, Waluchow provides a disclaimer at the outset that suggests a somewhat different approach to the question of the relationship between legal theory and metaethics:

I do not wish to become embroiled in any conflicts there might be as to the nature and objectivity of moral standards . . . . We will merely assume that people do appeal to standards like the principles of equality, liberty, fairness, and justice in assessing social institutions and their products; that these activities are not totally nonsensical as

30. See DWORKIN, TAKING RIGHTS SERIOUSLY, supra note 7, at 348.
some radical moral nihilists might argue, but are open to at least some
degree of rational argument and assessment; and that it is these kinds
of standards that we have in mind when we ask about the possible
role of political morality in determining the existence and content of
valid laws. [p. 2 n.3]

This passage helps clear up the question of Waluchow’s position
on the connection between legal theory and metaethics. Although
Waluchow is not always consistent,31 the best interpretation of his
view is that he himself does not take a stand on the question of the
status of moral judgments any more than Hart did. Rather, his
claim about the reality of moral judgments is a claim about how the
law sees itself, a claim about the view of judges and other officials
who are engaged in the practice of law. Legal systems that invoke
moral standards assume what moral philosophers endlessly debate,
namely, whether moral standards are sufficiently objective and ca-
pable of determinative answers to qualify as legal standards.

This interpretation — that Waluchow is describing and report-
ing on the implicit assumptions of law rather than defending a
metaethical claim of his own — is reinforced throughout the book.
At almost every critical juncture, Waluchow turns to the arguments
and opinions of judges to demonstrate his claims about how legal
systems invoke and apply moral standards. This interpretation
helps make Waluchow’s position consistent, but it has implications
for positivism that he fails to appreciate. I shall return to those im-
lications at the end of this review. First, it is time to look at the
second issue: How does Waluchow’s position differ from
Dworkin’s?

III. IS IT STILL POSITIVISM?

We have just noted that Waluchow’s method of argument relies
extensively on the evidence of judicial decisions — descriptions of
how judges invoke and respond to moral standards they are re-
quired to apply. This method bears an obvious resemblance to
Dworkin’s, which also relies on the way that judges write opinions
as evidence for claims about the connection between moral and
legal theory. The similarity in method, as well as the similar conclu-

31. Waluchow’s position may be that moral standards are (sometimes) “determinate” but
not necessarily “objective” and that only the former is necessary to qualify as law. In this
way, he could continue to remain detached from questions about the status of moral judg-
ments. But this position would also require much more in the way of defense and explora-
tion into metaethical questions than Waluchow provides. For example, if moral standards are
determinate only because of certain facts — facts about human beliefs or social conventions
— he will not have made a case for going beyond the exclusive positivist’s “social fact”
theory of law. The best interpretation of his position thus seems to be that he believes that
moral statements are (sometimes) determinate because of moral facts — not social facts —
which seems to be a commitment to some sort of claim about the objective status of such
judgments.
sion that moral standards can be legal standards, leads to the second major argument of the book: The incorporation of moral standards in law is consistent with positivism and distinguishable from Dworkin's version of natural law.

Two central ideas figure in this part of Waluchow's argument. First, Waluchow separates theories of adjudication from theories of law (p. 56). Dworkin finds the key to law in the rights litigants have before courts — moral rights to particular judicial decisions that are the consequence of both political morality and previous legal facts. Waluchow, in addition to suggesting technical difficulties with this account, has a "simpler" suggestion for explaining these rights. These rights, he argues, are not the result of a theory of adjudication that somehow has independent legal or moral status; rather, all such rights derive, in the end, from their source in a rule of recognition or similar "pedigree" of a positivist sort. After all, how judges decide cases is itself a matter of the instructions that judges receive, and those instructions can be altered by legislators in a manner consistent with positivism. Second, the pedigree that makes the theory positivist is a social fact — like Hart's rule of recognition — that can, but need not, identify moral standards as standards judges should use in deciding cases. We already have seen that this view distinguishes Waluchow's theory from that of the exclusive positivist, who insists that validity is determined by pedigree alone without reference to content, as would be required with moral standards. The same view of the kind of pedigree that determines legal validity distinguishes Waluchow's theory from natural law theories as well. The latter theories insist that some reference to content or morality is always necessary in determining legal validity; for Waluchow, it remains a contingent question which moral standards, if any, a particular legal system will choose to incorporate by reference in its basic pedigree.

32. Waluchow suggests that Dworkin's theory leads to the odd result that the law will be different depending on whether one is in a lower court, which cannot overrule a binding precedent, or a higher court, which can. See pp. 46-58. This result, he says, is inconsistent with Dworkin's insistence on personifying the legal community in a way that requires it to speak "with one consistent voice." See p. 53.

Waluchow's argument seems to overlook the possibility that the political principles that explain why lower and higher courts have different powers will themselves be part of the "law" and thus will explain "with one voice" why the results, but not the law, are different in different courts. Nothing in Dworkin's theory, in short, prevents his incorporating Waluchow's varying "institutional forces of law" as part of what courts must consider in deciding what legal rights citizens have.
IV. Is It Still Law?

A. The Limits of Pedigree

With one exception, the above account justifies Waluchow's claim that he has defended a theory of law that is at least as consistent and, arguably, as plausible as the main alternatives he considers — the exclusive positivism of Joseph Raz and the rights theory of Ronald Dworkin. But Waluchow's own argument contains an unexplored hint about the limits of even the best form of positivism. Waluchow's theory is primarily a pedigree theory of law. The pedigree has been modified to permit incorporation of moral standards, but the basic positivist claim — that law is a matter of source, with content serving as the test for validity only contingently as source permits — is the core of the theory.

Even if this is still positivism, it may not be all there is to law. The real debate, after all, is not over the concept of positivism but over the question of which theory of law best reflects our understanding of what law is. By Waluchow's own account, as we have seen, that question requires paying attention to how officials responsible for creating and applying legal norms view what they are doing. Waluchow finds that the practice of referring to moral standards reveals an implicit assumption of objective status for morality that allows for a contingent connection between law and morality. But if this is the test for law, then one must also provide some account for the other normative claims about the concept of law itself that the practice reveals. What are those normative claims, and how do they limit the ability of a pedigree theory to account fully for law?

Actually, it was positivism itself that first called attention to this normative posture of legal systems, and both of the positivists whom Waluchow examines most closely — Hart and Raz — are identified closely with this view, in contrast to the more classical coercive account of Austin and Bentham. Waluchow, presumably, does not advocate returning to a view of law as essentially coercive. If, however, law is essentially normative, then it must be true that a standard, in order to count as law, must do something besides guide conduct. It also must be capable of supporting certain normative claims that are made about it.

Most of Waluchow's consideration of this issue is confined to what he calls a "theory of compliance." In the same way that Waluchow separates theories of adjudication from theories of law, he also suggests that theories of compliance should be separated from theories of law. He argues that Dworkin invites confusion by trying to account for "wicked law" as law in one sense ("the fact of law") (p. 59), but not in another sense (it does not "morally license coercive force") (p. 61). This argument seems at odds with
Waluchow's own approach to the issue of metaethics. It may be confusing to the legal theorist to try to combine theories of compliance and theories of law, but the question is not whether separating these notions makes for a simpler or easier task for the legal philosopher. The question is rather whether this "confusing" combination of "law as fact" and "law as norm" is itself a reflection of how insiders, whose concept we are explicating, think of legal systems. Just as Waluchow avoids entangling himself in questions of metaethics by noting that judges and officials act as if moral standards are meaningful (however uncomfortable or odd that conclusion may be for the skeptical moral philosopher), so too he should consider more carefully what officials say about the normative nature of the concept of law itself: Law guides conduct, not in the way that coercive orders do, but rather in the way that normative judgments do, or purport to do.

B. Including Law's Normative Claims in Legal Theory

There are two reasons, I think, for Waluchow's failure to see that the normative claims made by law place limits on any pedigree-based account of legal validity. The first reason is his exclusive focus on morality in the sense of particular legal standards that courts contingently are invited to apply. By making this question central — what standards can and do courts apply in judicial decisions — he unduly slights the larger question concerning the moral connections insiders attach to the concept of law itself.

Second, when Waluchow does consider the suggestion that law may be a concept that refers both to the source of legal standards (the pedigree) and also to certain normative claims about those standards, he does not distinguish between two different kinds of normative claims that one might think are essential to law: claims about compliance and claims about the justification of coercion. Thus, in discussing Dworkin's mixing of theories of law and theories of compliance, Waluchow treats the claim that law justifies coercion as apparently the same as the claim that law justifies compliance. The difference between these claims, however, is important and worth a brief exploration.

1. Justifying Compliance

If only standards that actually created moral obligations to comply were to count as law, we would have a connection between legal and political theory of the strongest possible type. In fact, however, the question whether there is even a prima facie obligation to obey law remains controversial, and the suggestion that the obligation to obey law is absolute has few or no supporters. These features suggest that considering how insiders view the concept of law lends
some support to Waluchow's claim that we should and do separate theories of law and theories of compliance.33

2. Justifying Coercion

There is, however, a second claim about legal standards that cannot be separated from the concept of law so easily. When officials impose sanctions simply because pedigreed rules have been violated, they implicitly claim that the coercion is justified. This basic claim, which I have described elsewhere as the minimal claim that all legal systems make,34 differs from the claim that compliance is required in two significant respects. First, there is little disagreement that legal systems are justified in enacting and enforcing norms. The whole point of the state, after all, is to make decisions about what is to be done. Only an anarchist, who denies the state's right to exist in the first place, could fail to recognize that the state's right to interfere with a citizen's ability to do whatever he would otherwise be doing in a state of nature is a fairly easy first step for any political theory. It is only the separate and additional claim that citizens have duties to obey the state that is controversial. Second, this claim about the justification of state coercion attaches to law qua law — just because of its pedigree. It is a content-independent claim about the state's moral justification in acting, in good faith, on its own lights in determining how to govern society. These two features — the widespread acceptance of the minimal normative claim and the fact that the claim is connected to the concept or practice of law itself — make it more difficult to separate this moral claim from a theory of law.

3. The Essence of Law

To see why it is so hard to separate this moral claim about law from legal theory, let us see what happens if we try. What would be the consequence of claiming that we should identify law purely by its pedigree without concern for the separate question of whether standards so identified can serve as the basis for justified coercion? Why is it that this moral question cannot be separated from the fact of law in the same way that Waluchow separates theories of compliance from theories of law? To return to Waluchow's consideration of this issue in his critique of Dworkin, imagine a case of truly "wicked" law in which the conclusion is reached that no moral theory could justify enforcing the law. In other words, standards that appear to have the proper pedigree are nevertheless sufficiently un-

33. For the contrary view, see SOPER, supra note 17, at 101-07.
just that even insiders — officials charged with enforcing the law — agree that no plausible moral argument can justify coercing those who fail to comply. Can one still claim, nevertheless, that such standards are law simply because they continue to carry the same pedigree as other legal standards that are not so unjust? Because law is a practical concept, some practical conclusion, it seems, should follow the identification of a standard as "legal." Otherwise, as Dworkin notes, "we are suddenly in the peculiar world of legal essentialism." \(^{35}\) But the only practical conclusion that is left in the case of such wicked standards is the prediction of coercion or force — not justified force, but simply power to carry out the sovereign's will. Is this still law?

This ultimate question returns to the problem of the essence of a concept that always has haunted this field of jurisprudence. Without attempting a resolution of that question, one can note at least that the new direction in positivist thinking marked by Hans Kelsen and Hart — and carried through by Raz and other "modern" positivists — has been consistent in rejecting a purely coercive account of law as an adequate explication of the concept. Moreover, the force of these modern theories of positivism rests largely on the same evidence on which Waluchow bases his arguments for viewing moral standards as potential legal standards, namely the descriptive accounts of the practice of law revealed by the way that insiders use and refer to the concept of law. That we would be puzzled about what to call standards that have no moral consequence at all is some evidence that the moral qualification is not contingent but part of the essence of law. The case for natural law, it turns out, does not depend on complex theories of adjudication, à la Dworkin, nor does it depend on claiming that moral standards must always be among the standards courts use, along with tests of pedigree, to determine legal validity. The simplest case for natural law starts where Waluchow ends — essentially with a pedigree theory of law. It notes only that this pedigree, though it usually will be sufficient to determine legal validity, will fail as a test for law in certain extreme cases — extreme cases of the sort that bring to mind the Nuremberg principles and the gradual international recognition of moral limits on the power of positive law to have any practical effect other than that of pure coercion. Even if we start with positivism as the basic model and use pedigrees of either the exclusive or the inclusive type to identify law, we still must admit the qualification that if the law so identified is too unjust, it is not law and can be nothing but pure force.

I assume that Waluchow did not mean to return positivism to the classical view of Austin, but rather to continue a dialogue within the modern version of positivism that Raz and Hart endorse.\textsuperscript{36} If my assumption about Waluchow's intent is correct, his contribution serves as a welcome addition to the debate within modern positivism. But it also serves as a subtle, if unintended, reminder of the limits of pedigree theories of law.

\textsuperscript{36} Ultimately, the classical view may be the most consistent version of positivism and the inevitable position for those who insist on separating law as social fact from even the minimal normative claim insiders make about law.