WHAT IS THE RULE OF RECOGNITION (AND DOES IT EXIST)?

by

Scott J. Shapiro
WHAT IS THE RULE OF RECOGNITION (AND DOES IT EXIST)?
Scott J. Shapiro

One of the principal lessons of The Concept of Law is that legal systems are not only comprised of rules, but founded on them as well. In sharp contrast to Bentham and Austin who had insisted that the sovereign makes all of the rules, Hart argued instead that the rules make the sovereign. For as Hart painstakingly showed, we cannot account for the way in which we talk and think about the law – that is, as an institution which persists over time despite turnover of officials, imposes duties and confers powers, enjoys supremacy over other kinds of practices, resolves doubts and disagreements about what is to be done in a community and so on – without supposing that it is at bottom regulated by what he called the secondary rules of recognition, change and adjudication.

Given this incontrovertible demonstration that every legal system must contain rules constituting its foundation, it might seem puzzling that many philosophers have contested Hart’s view. In particular, they have objected to his claim that every legal system contains a rule of recognition. More surprisingly, these critiques span different jurisprudential schools. Positivists such as Joseph Raz, as well as natural lawyers such as Ronald Dworkin and John Finnis, have been among Hart’s most vocal critics.

In this essay, I would like to examine the opposition to the rule of recognition. What is objectionable about Hart’s doctrine? Why deny that every legal system necessarily contains a rule setting out the criteria of legal validity? And are these objections convincing? Does the rule of recognition actually exist?

This essay has five parts. In Part One, I try to state Hart’s doctrine of the rule of recognition with some precision. As we will see, his position on this crucial topic is often frustratingly unclear. Hart never tells us, for example, what kind of rule the rule of recognition is: is it a duty imposing or power conferring rule? Nor does he identify the rule of recognition’s audience: is it a rule practiced only by judges or by all legal officials? I also explore in this part whether the United States Constitution, or any of its provisions, can be considered the Hartian rule of recognition for the United States legal system.

In Part Two, I attempt to detail the many roles that the rule of recognition plays within Hart’s theory of law. In addition to the function that Hart explicitly assigned to it, namely, the resolution of normative uncertainty within a community, I argue that the rule of recognition, and the secondary rules more generally, also account for the law’s dexterity,

1  Professor of Law and Professor of Philosophy, Yale University. Thanks to Matt Adler, Daniel Halberstam and Scott Hershovitz for very helpful comments on an earlier version of this draft.

2 I once heard Jeremy Waldron describe Hart’s inversion of Austin in this way.
efficiency, normativity, continuity, persistence, supremacy, independence, identity, validity, content and existence.

In Part Three, I examine three important challenges to Hart’s doctrine of the rule of recognition. They are: 1) Hart’s rule of recognition is under- and over-inclusive, i.e., some rules that are part of a particular legal system are not so considered by his account and, conversely, some rules that his account deems to be part of a legal system are not in fact so; 2) Hart cannot explain how social practices are capable of generating rules that confer powers and impose duties and hence cannot account for the normativity of law; 3) Hart cannot explain how disagreements about the criteria of legal validity that occur within actual legal systems, such as in American law, are possible.

In Parts Four and Five, I address these various objections. I argue that although Hart’s particular account of the rule of recognition is flawed and should be rejected, a related notion can be fashioned and should be substituted in its place. The idea, roughly, is to treat the rule of recognition as a shared plan which sets out the constitutional order of a legal system. As I try to show, understanding the rule of recognition in this new way allows the legal positivist to overcome the challenges lodged against Hart’s version while still retaining the power of the original idea.

I. THE RULE OF RECOGNITION: AN INTRODUCTION

Hart formally introduced the rule of recognition in Chapter Five of The Concept of Law. There he considered a community which does not have a legal system and then invites the reader to ponder the various social problems that would arise in that group and how the introduction of certain rules (among them being the rule of recognition) would resolve these difficulties.

In a pre-law society, Hart supposed, all rules are customary ones. In other words, a rule exists within such a group if, but only if, it is accepted and practiced by most of its members. Hart then considered what would happen should some doubt or disagreement arise within the group about proper behavior. (Imagine that some members believe that a person should be able to take two mates, whereas others think that the limit should be one.) Since the only property that their rules share is their acceptance by the group, there will be no other common mark to which members can point (e.g., an inscription in some authoritative text, declaration by some official, etc.) in order to resolve their controversy.

5 Id. at 92.
Hart claimed that this normative uncertainty would be unproblematic in a small group united by bonds of kinship and inhabiting a stable ecological niche. Presumably, relatively few doubts and disagreements would arise in such groups and those that did could be overcome through either head counting to determine the existing custom or some combination of persuasion, deliberation and negotiation. However, as groups expand and become more heterogeneous, or when environmental conditions are highly fluid, uncertainty will likely proliferate and these techniques will become more costly or less effective. And given that the need for dispute resolution is bound to be great within such a group, the insecurity engendered by these doubts and disagreements will be distressing, perhaps even crippling.

Normative uncertainty is not the only problem facing such groups; customary rules also possess a “static character” that renders them defective tools for regulating all but the smallest human communities. Suppose there is sudden need for the group to act in a certain manner, e.g., to increase the amount of grain that each family contributes to communal storage as a result of drought. The simplest and quickest response would be for some members of the group to deliberately change the rules, e.g., to amend the tithing rules. However, in a group governed solely by custom, this option is unavailable. Their rules cannot be changed at will: customary rules vary only through a slow process of growth and decay. The urgent need of the group to respond to the drought, therefore, will likely go unmet.

Finally, Hart considered the “inefficiencies” associated with this simple regime of customary rules. Suppose there is a clear rule about how land is to be acquired. It is accepted custom, say, that the first person to stake his claim is the rightful owner. What happens, though, when there is factual disagreement about who is the first-claimant? Since the regime contains no mechanism for determining the satisfaction or violation of any of the rules, the attempt to settle who actually staked the claim first will likely be costly and could even turn ugly.

Hart suggested that the fundamental rules of legal systems solve the various defects of pre-legal, customary societies. Legal systems address the problem of uncertainty by providing a rule which determines which rules are binding. By referring to this rule about rules – what Hart termed the “rule of recognition” – normative questions can be resolved without engaging in deliberation, negotiation or persuasion. If there is a doubt about, say, how many mates are acceptable, the rule of recognition can direct the parties to the authoritative list of rules on the rock in the town square, the past pronouncements of the village elder, the practice of other villages and so on, to determine the answer.

---

6 Id.
7 Id. at 92-93.
8 Id. at 93-94
9 Id. at 94.
The static character of customary norms is overcome by what Hart
called a “rule of change.” A rule of change confers power on a person or
institution to create, modify or extinguish rules and may also specify the
procedures to be used in exercising that power. Since the rule of change
empowers certain persons or bodies to amend the rules, behavior may be
shifted in the desired direction through the exercise of legal authority. A
group facing a drought can, for example, deliberately change the tithing
rules and hence address the dire circumstances in an expeditious manner.

Finally, the problem of inefficiency is solved by what Hart called a
“rule of adjudication.” This rule confers the power on certain bodies to
apply the rules, i.e., to determine whether a rule has been satisfied or
violated on a particular occasion, and specifies the method to be followed
in adjudication. In our example of first claimants, the body identified as
the authoritative adjudicator would have the power to determine which
person first claimed the land and hence who is the rightful owner of the
property.

We are now in a position to state Hart’s doctrine of the rule of
recognition in a more abstract manner. According to Hart, every legal
system necessarily contains one, and only one, rule which sets out the test
of validity for that system. The systemic test of validity specifies those
properties the possession of which by a rule renders it binding in that
system. Any norm that bears one of the marks of authority set out in the
rule of recognition is a law of that system and officials are required to
recognize it when carrying out their official duties.

In the course of setting out the criteria of legal validity, the rule of
recognition also specifies orders of precedence among sources of law. In
the United States, for example, the rule of recognition mandates that
federal law trumps state law, federal constitutional law trumps federal
statutory law and constitutional amendments made in accordance with
Article V trump earlier constitutional provisions. Hart called tests such
as the one set out in Article V “supreme” criteria of legal validity, because
they specify those legal rules which are not trumped by any other possible
rule.

The most salient property of the rule of recognition is that it is a
secondary rule. It is a rule about the validity of other rules (i.e., the
“primary” rules). The rule of recognition is also a social rule. It is
“social” in two different senses. First, the rule of recognition exists and
has the content it does because, and only because, of certain social facts.
In particular, its existence and content is determined by the fact that
members of a group take the internal point of view towards a standard of
conduct and use it to evaluate the validity of norms and the behavior that

---

10 Id. at 95.
11 Id. at 97.
12 See, e.g., U.S. CONST. art VI, § 2 and U.S. CONST. art V.
13 Id. at 106.
14 Id. at 110.
falls within their purview.\(^{15}\) Second, the rule of recognition is social in the sense that it sets out a group-wide standard. Members of this group do not accept this rule “for their part only,” but rather treat the standard it sets out as the official way in which the law is to be determined in their community.\(^ {16}\)

Because the rule of recognition is a social rule, it is capable of being an ultimate rule.\(^ {17}\) It is ultimate in the sense that it does not exist in virtue of any other rule. Its existence is secured simply because of its acceptance and practice. The primary rules of the legal system, by contrast, are not ultimate because they exist in virtue of the rule of recognition. The rule of recognition validates, but is not itself validated.

Some Complications

Stating the basic idea behind Hart’s doctrine of the rule of recognition is easy enough; formulating the doctrine with greater precision, however, is surprisingly difficult. For example, what is the basic form of the rule of recognition? Astonishingly, Hart was vague on this critical point. Hart often characterized the rule of recognition as a test of what the law is in a particular legal system.\(^ {18}\) Thus, he described the British rule of recognition as “whatever the Queen in Parliament enacts is law.”\(^ {19}\) On this interpretation, then, the rule of recognition has the following canonical form: “Any norm that bears properties \((A_1, \ldots, A_m)\), \((B_1, \ldots, B_n)\) or \((C_1, \ldots, C_o)\) is a law of system S.”

Treating the rule of recognition simply as a test of legality, however, fits uncomfortably within the Hartian framework which famously acknowledges only two types of legal rules, namely, duty-imposing and power-conferring.\(^ {20}\) On their face at least, tests are neither. The scientific criterion which states that a substance is acidic if its pH is higher than 7 and the linguistic test which defines a “bachelor” as an unmarried male does not confer a power or impose a duty. Is it possible, then, to understand the rule of recognition either as a power-conferring or duty-imposing rule?

I think that the first option cannot be Hart’s position. For if we suppose that the rule of recognition in Britain is: “The Queen in Parliament has the power to create British law,” we inadvertently convert Britain’s rule of recognition into its rule of change. Moreover, the rule of recognition can validate certain types of customs, and since customs need

---

\(^{15}\) On Hart’s account of the internal point of view, see Scott J. Shapiro, *What is the Internal Point of View?*, 75 FORDHAM L. REV. 1157 (2007).

\(^{16}\) Id. at 115-16.

\(^{17}\) Id. at 107-08.

\(^{18}\) See, e.g., id. at

\(^{19}\) Id. at 107. See also id. at 68: “In the simple society of Rex it may be the accepted rule … that no law of Rex shall be valid if it excludes native inhabitants form the territory …”

\(^{20}\) See id., ch. 3. See also JOSEPH RAZ, *THE AUTHORITY OF LAW* 92 (1979).
not be (and usually are not) created through the exercise of legal authority, the rule that validates them cannot be power-conferring.

The only alternative, then, is to treat the rule of recognition as a duty-imposing rule. The rule of recognition, on this account, imposes a duty on officials to apply rules that bear certain characteristics. In our British example, it requires members of the British legal system to apply the rules enacted by the Queen in Parliament. In the United States, the rule of recognition requires, at least in part, all federal and state officials to apply those rules which regulate interstate commerce, have been enacted by a majority of both houses of Congress and signed by the President (or a super-majority of both houses after veto by the President).

This interpretation of Hart’s doctrine, however, might raise the following concern: why does Hart present the rule of recognition as a test when in reality it is a duty imposing norm? The answer, I believe, is that, according to Hart, the law consists of all the norms that legal participants are under a duty to apply in their official capacities. In other words, the rule of recognition sets out the criteria of legal validity, and hence picks out the set of legal rules for a particular legal system, because the law of a particular system just is the set of rules that officials of a certain system are under a duty to apply and the rule of recognition sets out the content of this duty.

If this interpretation of Hart’s doctrine is correct, it follows that the vast majority of the text of the United States Constitution does not set out the United States rule of recognition. Article I, Section 8, for example, begins: “The Congress shall have Power To lay and collect Taxes, Duties, Imposts and Excises, ….” This provision formulates part of the federal rule of change insofar as it confers power on Congress, not imposes a duty on officials. Article V and VII are also part of the rule of change, for both provisions confer power on state legislatures and conventions to ratify and amend the Constitution and specify the procedures to be used. Similarly, most of Article III is best understood as part of the federal rule of adjudication, for it confers power on the Supreme Court, and on any lower federal courts that Congress should happen to create, to decide certain cases, as well as partially specifying the method that courts should follow when engaged in adjudication.

If the rule of recognition imposes duties on legal officials, we might wonder which officials. Sometimes, Hart made it appear as though the rule of recognition applies to all officials. He wrote: “There are therefore two minimum conditions necessary and sufficient for the existence of a legal system. On the one hand, those rules of behavior which are valid according to the system’s ultimate criteria of validity must be generally

---

22 For a similar interpretation, see Raz, supra note 20, at id.
23 U.S. CONST. art I, § 8.
obeyed, and, on the other hand, its rules of recognition specifying the
criteria of legal validity and its rules of change and adjudication must be
effectively accepted as common public standards of official behavior by its 
oficials. Other times, Hart focused exclusively on judges. “[I]t is
the case that this rule of recognition … is not only accepted by him but is
the rule of recognition actually accepted and employed in the general
operation of the system. If the truth of this presupposition were doubted,

24 Id. at 115-16 (emphasis added).
25 Id. at 108 (emphasis added). See also Hart’s description of rule of recognition in the
Postscript as a “judicial customary rule” at 256.
26 The Supremacy Clause of Article VI, § 2 does impose such a duty on state judges.
27 U.S. CONST. art VI, § 3.
28 See, e.g., Glanville Williams, ed., SALMOND ON JURISPRUDENCE 41 (Eleventh Edition,
1957) (“The law consists of the rules recognized and acted on by courts of justice.”)
quoted in Raz, supra note 21 at 190.
II. THE ROLES OF THE SECONDARY RULES

To be sure, there are many more questions we could ask about the particularities of Hart’s doctrine; indeed, much of recent Anglo-American legal philosophy has concerned itself with debating the exact nature of the rule of recognition. Fortunately, these details need not concern us. For now at least, we should have a firm enough grasp of Hart’s theory to be able to understand the challenges to it and assess their cogency.

Before I go on to examine these various objections, however, I would like to spend a bit more time exploring the various roles that Hart thought the rule of recognition, and the secondary rules more generally, play in a legal system. As I hope to show, the rules of recognition, change and adjudication are absolutely indispensable for making sense of a whole range of legal phenomena. This being the case, our puzzle will deepen: how can anyone sensibly reject Hart’s doctrine and deny the existence of the rule of recognition?

Resolution of Normative Uncertainty, Dexterity and Efficiency

In Part One, we saw one role that the rule of recognition plays in all legal systems, namely, the resolution of normative uncertainty. According to Hart, the rule of recognition resolves doubts and disagreements within a group about which primary rules to follow. It does this by picking out properties of primary rules the possession of which mark them as binding.

We also saw that the rule of change advances the dexterity of the law. When in place, the law has the ability to adapt nimbly to changed circumstances. Those designated by the rule of change need not wait for custom to evolve; rather, they have the power to deliberately alter the rules and thus enable the group to meet the urgent challenges that they face.

Finally, the rule of adjudication promotes the efficiency of the law. In a group fortunate to contain such a rule, disputes concerning the satisfaction or violation of a norm need not drag on and ripen into feuds. When an empowered adjudicator determines that a rule has been broken,

---


31 The rule of recognition, of course, does not resolve all normative uncertainty. It only resolves those doubts and disagreements which arise in cases regulated by the norms that it happens to recognize.
this decision is supposed to settle the disagreement. The judgment is authoritative and is to be supported by the social pressure that law typically brings to bear.

Continuity, Persistence and Normativity

According to Austin, legal sovereignty is created by asymmetrical habits of obedience: the sovereign is the one who is habitually obeyed by the bulk of the population and who habitually obeys no one else. Hart effectively showed that habits of obedience cannot create sovereignty. First, habits are not “normative,” i.e., they are incapable of generating rights or obligations all by themselves. Second, habits cannot establish the “continuity” of legal authority: Rex I’s successor, Rex II, will be the sovereign from the moment he takes office even though Rex II has yet to be the object of habitual obedience. Third, habits cannot establish the “persistence” of law: Rex I’s laws will be legally valid after his death despite the fact that the dead cannot be habitually obeyed.

Instead, Hart argued, sovereignty is created by rules, not habits. Rules are normative: they are capable of conferring rights and imposing duties. Moreover, rules can account for the continuity of legal authority: Rex II has the power to legislate from the moment of Rex I’s death because the legal system contains a secondary rule of change giving him the power to do so. Finally, rules can explain the persistence of law: Rex I’s rules are valid even after his death because the rule of recognition requires judges to apply all the rules made by past kings.

Supremacy and Independence

In addition to resolving normative uncertainty and accounting for the dexterity, efficiency, normativity, continuity and persistence of law, Hart also showed that the secondary rules can be used to explain two properties shared by modern state legal systems: supremacy within its borders and independence from other systems. In contrast to Austin’s account according to which these properties arise from the asymmetry in habitual obedience, i.e., Rex is supreme because he is habitually obeyed and habitually obeys no one else and his regime is independent because he habitually obeys no one else, Hart credited the rule of recognition. In Rex’s kingdom, the rule of recognition requires all officials to privilege Rex’s will and hence rendering his power supreme above all others. The independence of Rex’s legal system is established in a similar manner. Since the rule of recognition refers to Rex’s enactments, those of his subordinates and no one else’s, the system formed will have a separate

---

32 Hart, supra note 3, at 60.
33 Id. at 53.
34 Id. at 62.
35 Id. at 106.
existence from all other legal regimes (which have their own rules of recognition).36

At the same time, Hart’s account of sovereignty does not imply that the sovereign is necessarily “above the law.”37 In a constitutional regime, the secondary rules will typically limit the supreme and independent powers of the sovereign. Although the American people are sovereign in the United States and have the power to amend the Constitution, the Constitution nonetheless limits their power to do so, both by making certain provisions unalterable and prescribing an extremely onerous procedure that must be followed before an amendment is ratified.38

Identity

Hart also pointed out that secondary rules are necessary to distinguish legal systems from other collections of norms, such as games, religions, corporations, clubs, etiquettes, popular moralities, …, etc. According to Hart’s famous dictum, law is best understood as “the union of primary and secondary rules.”39 Thus, a legal system differs from etiquette because the latter consists solely of primary rules, whereas the former also contains rules about these rules. Hart does not claim, of course, that the union of primary and secondary rules completely distinguishes legal systems from all other normative systems. The rules of corporations, for example, contain secondary rules as well. There are rules about who can change the rules of the corporation and which rules corporate officers are required to recognize when doing their job. Yet, corporations are not legal systems. The postulation of secondary rules is at best only partially constitutive of the identity of law.

Validity, Content and Existence

In contrast to pre-legal societies, which according to Hart are governed purely by custom, legal systems can, and typically do, contain some rules that are not themselves practiced by members of the group. Jaywalking, for example, is prohibited in New York City even though most everyone does it. On Hart’s account, the rule prohibiting jaywalking exists because it is validated by the New York City rule of recognition that requires legal officials to heed rules enacted in similar fashions.

For Hart, then, the rule of recognition secures the existence of all primary rules. As long as a rule bears the characteristics of legality set out in the rule of recognition, it exists and is legally valid. Indeed, Hart claimed that the concept of validity is used precisely in those contexts

36 Id. at
37 Id.
38 See, e.g., U.S. CONST. art V.
39 This phrase is the title of Chapter 5 of THE CONCEPT OF LAW.
where the existence of rules does not depend on their being practiced. To say that a rule is valid is to express a judgment that it is binding because it passes the test of some other existing rule, and not because it is accepted by its audience from the internal point of view.

Aside from establishing the validity of all the primary legal rules, the rule of recognition determines the membership, or content, of particular legal systems. On Hart’s account, the rule of recognition of S determines all and only the laws of S. Thus, the New York State Statute of Fraud is not simply binding according to New York law—it is part of New York law.

Finally, the rule of recognition secures the existence of legal systems. According to Hart, a legal system exists for a group G just in case (1) the bulk of G obeys the primary rules and (2) officials of G accept the secondary rules of recognition, change and adjudication from the internal point of view and follow them in most cases. Thus, even if it turned out that most of the citizens of Rhode Island obey most of the rules of Roman law, it would not be true that Roman law still exists today, given that the Rhode Island State officials would not be following the secondary rules of the (extinct) Roman legal system.

III. THREE OBJECTIONS

Having set out Hart’s doctrine of the rule of recognition, I would like to rehearse three important objections that philosophers have lodged against it. The first challenge concerns Hart’s claim that his account accurately characterizes the content of a legal system, while the second and third relate to his claim that the rule of recognition is necessarily a social norm. As we will see, these objections do not challenge Hart’s general thesis that the law rests on secondary rules. Rather, they seek to undercut his specific claims about the nature of these rules. Whether these challenges are successful will be taken up in the last two parts of the paper.

First Objection: Under- and Over-Inclusiveness

Any theory that purports to characterize a legal system’s content must ensure that, for every system, it specifies all and only those norms that belong to that system. The theory will fall short, therefore, to the extent that it is either under- or over-inclusive. The first objection is that Hart’s theory fails in both these respects.

Recall that on Hart’s theory, the content of a legal system is established by that system’s rule of recognition. The New York Statute of

---

40 Id. at 108-10.
41 Id. at 95, 103.
42 Id. at 113-17.
Frauds is part of New York law, and not, say, New Jersey law, because the statute is valid according to the New York, and not the New Jersey, rule of recognition.

It is important to see that on Hart’s account the rule of recognition can characterize the content of a legal system only because it is one rule. Suppose, for example, the Governor of New York issues an executive ruling. Hart would say that this executive order is part of New York law because it is endorsed by the same rule of recognition that validates the Statute of Frauds. The unity of New York law, therefore, is secured by the unity of New York’s rule of recognition.

As John Finnis and Joseph Raz have objected, however, Hart does not explain what makes the rule of recognition a rule, as opposed to rules, of recognition. Why think that the rule that validates executive orders of the governor is the same one that validates the regulations enacted by the New York State legislature? Hart, it seems, is able to establish the content of the law only by helping himself to the oneness of the rule of recognition. But without establishing the unity of the New York rule of recognition, he cannot show why the Governor’s orders ought to be included within the set of New York law. Indeed, on Hart’s own theory of rule-individuation, according to which rules which guide different audiences ought to be considered separate rules, many of the provisions of a Hartian rule of recognition do not properly belong to the same rule. In any complex system, different officials will be under duties to apply different rules. When this is so, there will be multiple rules of recognition and hence the rules that they validate will not be part of the same legal system.

Hart’s theory is not only under-inclusive, but over-inclusive as well. For it can easily be shown that Hart’s rule of recognition transforms the law into a vortex that sucks the rules of other normative system into its voracious maw. As Joseph Raz argued, judges are often under an obligation to apply laws of other jurisdictions in conflict of law cases. Hart, supra note 3 at 38-42.


44 See Hart, supra note 3 at 38-42.

45 At one point, Hart sought to establish the unity of the rule of recognition by claiming that any rule of recognition that sets out multiple criteria of legal validity will also contain a provision determining the order of precedence in cases of conflict. “The reason for still speaking of ‘a rule’ at this point is that, notwithstanding their multiplicity, these distinct criteria are unified by their hierarchical arrangement.” H.L.A. Hart, Book Review, 78 Harv. L. Rev. 1281, 1293 (1965) (reviewing Lon Fuller, The Morality of Law (1964)). I must confess to not understanding Hart’s argument. Why does the mere fact that a rule ranks certain criteria sufficient to incorporate those criteria into the rule? Furthermore, it is not clear how Hart would explain the unity of a rule of recognition which set out multiple sources of law but did not contain a conflict resolution provision.

defendant, a New York court may be required to apply the New Jersey Statute of Frauds. On Hart’s treatment of the rule of recognition as a duty-imposing rule, however, the New Jersey law would automatically become incorporated into New York’s law because judges would be under a legal obligation to apply it in certain cases. But this is clearly wrong: New York law does not annex New Jersey law simply because there are occasions when New York officials are required to apply the rules adopted by New Jersey officials.

Second Objection: Social Rules are Normatively Inert

Hart criticized Austin’s theory of sovereignty by pointing out that habits are not “normative,” i.e., that they are incapable of conferring rights and imposing duties. The legal power of the sovereign, therefore, cannot be explained simply by noting that others are in a habit of obeying him or her. But as Ronald Dworkin pointed out, it is unclear how Hart’s theory dodges the same bullet. After all, the secondary rules of a legal system exist if, but only if, they are accepted and practiced from the internal point of view. The rule of recognition, for example, need not be morally acceptable – it need only be followed. But how does the mere fact that certain judges think that they should follow certain rules, and act on this judgment, make it the case that any other judge ought to do so as well? If mere habits cannot impose duties and confer powers neither can mere practices.

According to Dworkin, Hart’s account conflates a “social” with a “normative” rule. When we assert the existence of a social rule, Dworkin claims, we are simply indicating that most members of the group accept the rule. In merely recognizing the practice, we are not thereby endorsing it. A group, for example, may seriously frown on inter-racial marriage and we may describe this racist practice by stating that in this group there is a (social) rule against miscegenation.

A “normative” rule, on the other hand, necessarily provides reasons for action. If we criticize someone for violating the rule against smoking indoors, we are not simply asserting that most others do not smoke indoors and would criticize others for doing so. Rather, we are identifying the ground of our criticism: smoking indoors is wrong because there is a (normative) rule against it.

In order to account for the sovereign’s right to rule and the judicial duty to apply the law, Dworkin concluded, it is not enough to postulate the existence of social rules. The mere fact that judges treat certain rules as valid is not dispositive as to whether they ought to do so. Only normative rules are normative – only they can confer rights and impose duties.

This second challenge to Hart’s doctrine, therefore, does not deny that there are secondary legal rules which impose duties and confer

---

48 Id. at 50-51.
powers. Rather, it asserts that these rules are not social in nature. For if secondary rules are to have normative power, they cannot exist simply because they are accepted from the internal point of view and followed in most instances. On this view, secondary rules can confer rights and impose duties only when they are also grounded in moral facts, namely, those that create a moral duty for judges to apply certain rules and confer moral legitimacy on persons to change and apply those rules.

Third Objection: The Incoherence or Insincerity of Disagreements about the Ultimate Criteria of Validity

On Hart’s theory, the rules of recognition, change and adjudication derive their content solely from consensus. The rule of recognition in the United States, for example, validates rules enacted by Congress, signed by the President and which regulate interstate commerce because most judges/officials take the internal point of view towards such a test.

Yet, as Ronald Dworkin has famously argued, this account of the criteria of legal validity is seriously flawed. For if it were correct, widespread disagreements about such criteria would be highly problematic. Since the criteria of legality are supposedly fixed by consensus, any pervasive disagreement about their content would indicate the absence of consensus, and hence the absence of a fact of the matter over which disagreement could be had.

The most obvious examples of disagreements over the criteria of legal validity are disputes about interpretive methodology. Many judges, for example, think that the proper way to interpret provisions of the United States Constitution is via the meaning that the public would associate with the provision at the time of its ratification (a view often called “public meaning originalism”). Others believe that constitutional provisions should be interpreted in light of current social mores, even if these attitudes contradict the original public meaning of the provision in question (a view often called “living constitutionalism”).

As Dworkin correctly points out, the dispute over originalism is best understood as a dispute about the criteria of legal validity. Originalists believe that the criteria of legal validity in the United States are originalist in nature: a rule of constitutional law is valid only if it corresponds to the original public meaning of a constitutional provision. Living constitutionalists, on the other hand, deny this characterization. Moreover, this disagreement over interpretive methodology is both prevalent and common knowledge: everyone knows that this disagreement is widespread.


50 Dworkin, supra note 49 at 29-30. Although Dworkin does not talk about “criteria of legal validity,” but instead about the “grounds of law,” the former can be defined in terms of the latter. On this point, see Shapiro, supra note ? at 40-41.
and everyone knows that everyone knows that this disagreement is widespread.

However, if Hart is right about the rule of recognition, the disputants are either insincere or incoherent. If judges do not in general agree about the correct way to interpret a constitutional provision, there can be no correct way to do so. It follows that taking a position on such interpretive matters amounts to political chicanery, confused thinking or both.\(^{51}\)

Dworkin concludes that the criteria of legal validity are determined not by social facts alone, but by moral facts as well. The virtue of such a position is that it can establish the possibility of such fundamental disputes: disagreements about the criteria of validity, on Dworkin’s view, reflect the fact that officials disagree about the moral value of law and/or its relation to practice.

**IV. SHARED PLANS**

This paper began with a puzzle, namely, how any theorist could object to Hart’s doctrine of the rule of recognition. The solution I hope should now be apparent. Critics of Hart’s doctrine do not deny that the law is either founded on rules or that the notions of legal authority and obligation are rule-based concepts. No one is proposing a return to Austin. Rather, these objections focus on the specific nature of the fundamental legal rules. Critics deny that there is a rule of recognition in Hart’s particular sense, which is to say a unitary duty-imposing norm

---

\(^{51}\) In his defense of Hart, Brian Leiter accepts this very conclusion. See Brian Leiter, “Explaining Theoretical Disagreements,” available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1004768. According to Leiter, the test of a theory is how well it fits the *totality* of the data. If Hart’s account does a better job than Dworkin’s in accounting for the whole range of legal phenomena, as Leiter believes it does, we must conclude that disagreements about interpretive methodology are indeed either insincere or confused. Leiter’s methodological point is clearly correct: no jurisprudential theory can be expected to validate *every* intuition that lawyers have about the practice in which they engage. Yet, I think that Leiter underestimates the theoretical importance of this data point. The idea that the criteria of legality are determined by consensus is not just one aspect of the practice among many; on Hart’s account, it is the fundamental ground rule of law. What ultimately makes it the case that some rule is a binding legal rule is that it is validated by some standard accepted by officials of the group. And herein lay the problem for Hart: the prevalence of disagreements about the criteria of legality, and the complete absence of criticism for engaging in them, strongly suggests that competent legal practitioners do not follow the ground rules that Hart claims they do. To be sure, this evidence is not dispositive. It is possible that legal experts are so confused about the practice in which they are engaged that they are simultaneously committed to mutually incompatible sets of fundamental ground rules. Sometimes they act on Hartian ones; other times they act on Dworkinian ones. Yet, as a methodological matter, any theory which flouts the principle of charity so brazenly should be severely penalized. The importance of Dworkin’s challenge, therefore, is that it reveals the heavy theoretical costs incurred by Hart’s version of legal positivism (as opposed to demonstrating logically that such a model must be false).
which sets out the criteria of validity whose existence and content derive from consensual practice among legal officials.

In the remainder of this essay, I would like to respond to these objections. In this Part, I will suggest that the key to answering them involves re-conceiving the secondary rules of a legal system as elements of a much larger shared plan which sets out the constitutional order of a legal system. The function of this plan is to guide and organize the behavior of legal officials through the specification of roles that each is to play in the collective activity of legal regulation. I will then argue that the rule of recognition should be identified with all of the norm-creating and – applying parts of this shared plan. These provisions determine the content of a particular legal system, as well as play all of the other roles that Hart ascribed to his version of the rule recognition. In Part Five, I try to show how this re-conceptualization helps to resolve the above objections.

Second-Order Uncertainty

To motivate my account of the secondary rules as constituting the major elements of a shared plan, I would like to return to Hart’s creation myth which was set out at the beginning of Part One. In his recounting, Hart dwells on the doubts and disagreements that arise in pre-legal communities concerning the obligations of private parties. Call this “first-order” uncertainty. According to Hart, the rule of recognition is needed to resolve these sorts of doubts and disagreements, which it accomplishes by picking out the primary rules that the group is obligated to follow.

We can imagine another type of uncertainty, one which does not concern private behavior, but rather the legitimacy of public officials to settle first-order uncertainty. Call this “second-order” uncertainty. In Hartian pre-legal communities, it is highly likely that second-order doubts will be as common as first-order ones. Just as group members can be uncertain as to whether a person is permitted more than one mate, they can have doubts and disagreements about whether, say, Rex gets to answer that question. Some members of the community, for example, might object to this proposed royal allocation of power and insist that the will of the majority be respected on these sorts of issues; the aristocracy, on the other hand, might be inclined to trust such power to one of their own. Still others might think that choosing mates is an individual and inviolable right that even democratic majorities cannot eliminate.

In such groups, the most obvious source of second-order uncertainty will be differing views about political morality. Since many people disagree about the natures of justice, equality, liberty, privacy, security and the like, they are bound to disagree about the proper form that government ought to take.

52 See, e.g., Hart, supra note 3, at 91 (“the rules must contain in some form restrictions on the free use of violence, theft and deception …”)
But there is another reason, often overlooked by legal theorists, for why issues of institutional design are bound to lead to normative uncertainty. Political questions about who should have power and how they should exercise it are intimately connected to questions of trust. Legal systems are constituted by delegations of awesome power to individuals – power that can be, and often has been, exploited to devastating effect. Conferring authority on those of ill will not only endangers mundane political objectives but more importantly, and ominously, provides a fertile environment in which tyranny and anarchy can grow. The need to discriminate between the trustworthy and the untrustworthy, therefore, will always be a central and pressing concern of legal design.

Because proper institutional design ought to track correct judgments of competence and character, disagreements about the latter will induce disputes about the former. And disagreements about trust are likely to arise within political communities because questions of who is trustworthy to do what, like issues of political morality, are highly complex and contentious.

Settlement

As I have argued, Hart neglected to recognize an important type of normative uncertainty that would take hold in a pre-legal community. In groups not linked by bonds of kinship, belief or value, doubts and disagreements would not only arise between the members as to what is to be done but also as to who has the authority to resolve these sorts of questions.

Recognizing the prevalence of second-order, as well as first order, uncertainty is imperative, for the resolution of the latter cannot be had without the resolution of the former. In other words, public officials can resolve the doubts of, and disagreements between, private parties only if members of the group are not uncertain about the identity of the public officials. If Rex intends for everyone to increase the amount of grain tithed, the group must know that they are supposed to listen to Rex before that intention can be fulfilled.

Thus, if a legal system is to resolve first-order questions about what private parties should do, it must be able to settle second-order questions first. But like the resolution of first-order uncertainty, settling complex and contentious questions of institutional design on an improvised, ad hoc basis, or through the forging of communal consensus, will likely be unachievable, or attainable only at a prohibitive cost. Moreover, even when questions of political power are not based on such complex and contentious issues of moral principle and social psychology, they are often generated by massive coordination problems which defy spontaneous or consensual solutions.
Legal systems are able to function effectively, I would like to suggest, because they resolve questions relating to the proper moral goals of the system, the competence and goodwill of legal actors, and how to coordinate behavior in pursuit of the proper goals via its secondary rules. In particular, some rules settle the content and contours of official duty, whereas others determine the scope of legislative, judicial and executive powers. These secondary rules resolve second-order uncertainty in an economical fashion. Instead of requiring members of the community to deliberate, negotiate, bargain or simply guess about the proper distribution of political power, they can appeal to the secondary rules of the system in order to resolve some of their doubts and disagreements.

Insofar as the task of the secondary rules is to determine the roles that legal officials of a particular system are to play, we might see them as constituting parts of a much larger plan shared by those officials. The constitutional law of a system, in other words, represents a plan for governance. Like all plans that regulate collective activities, the function of this shared plan is to guide and organize the shared activity of legal officials. It seeks to overcome the enormous complexity, contentiousness and arbitrariness associated with arranging a system of social regulation. Because reasonable (and unreasonable) people can have doubts and disagreements about which social problems to pursue and who should be trusted to pursue them, it is essential to have a mechanism that can settle such questions, creating a mesh between legal officials and channeling them all in the same direction.

The shared plan of a legal system, therefore, must settle questions of political morality by determining which goals and values a particular system should pursue and realize. It must determine whether and when equality trumps efficiency, security trumps privacy, the minority trumps the majority, faith trumps science, tradition trumps innovation and so on. These choices are normally manifested in the constitutional order, such as when a system that prizes democratic participation makes provisions for voting, representation, elections and some protection for public deliberation or a theocratic one empowers clerics to decide matters of principle and policy and minimizes the degree to which secular forces can affect the direction of the law.

Likewise, the shared plan of a legal system must allocate power and authority on the basis of the certain judgments of competence and character. Indeed, different constitutional configurations normally reflect these differing assessments of trustworthiness. Individuals who are judged to be less trustworthy are accorded fewer powers and subject to greater scrutiny than those who are judged more dependable. Because power normally tracks trust, it is useful for many purposes to conceptualize the distribution of rights in a legal system as a distribution of trust, or as I will call it an “economy of trust.” Monarchies, for example, can be understood as based on radically egalitarian economies of trust, where only royalty is trusted to set the terms of social cooperation. By contrast, democracies
are based on more egalitarian economies, where trust is widely distributed to its citizens. Systems of absolute legislative supremacy dole out greater trust to legislators than ones with judicial review. Regimes with unitary executives distrust committees to make decisions and hence grant a monopoly of trust to one person, whereas those with plural executives are more suspicious of individuals with large concentrations of power and hence disperse trust over a greater number of persons.

To say that a legal system’s shared plan resolves second-order normative uncertainty is not to claim, of course, that it resolves all such uncertainty. Plans, as Michael Bratman has emphasized, are typically partial: they settle certain questions about what is to be done, but leave other issues undecided. My initial decision to go to Mexico for vacation settles the general issue of destination but not the specifics of the journey. Plans are meant to be filled in over time as the future becomes clearer and the time for action approaches. Similarly, a constitution might confer the right to free speech, thus establishing that there is such a right, without setting its exact scope, weight or content. These questions are typically delegated to other bodies, such as courts, to decide. Constitutional adjudication, therefore, should be understood as a form of social planning, where the system’s shared plan is filled in over time and thus rendered more complete and informative.

Sharing a Plan

According to what I will call the “planning theory of law,” legal activity is best seen as structured by a shared plan. The function of this plan is not only to resolve first-order uncertainty about the obligations of private parties, but also to resolve second-order uncertainty about the rights and responsibilities of legal officials. The secondary rules of a legal system are thus seen as constituents of this shared plan, imposing and conferring law-creating and -applying duties and powers.

In claiming that officials of a particular legal system always share a plan of governance, I have been tacitly presupposing an account of plan-sharing. In the interests of full and fair disclosure, therefore, let me briefly sketch out such an account. What must obtain before we can say that a group shares a plan? On the account that I favor, a group shares a plan when the plan was designed, at least in part, for the group so that they may engage in some joint activity and the members commit to doing their parts and not to interfere with the others doing their parts. My friend and

---

54 On the importance of the qualification that shared plans need only be designed “in part” for the group that shares it, see note 56 infra.
55 Because a plan that is completely secret cannot be shared, we should also add that a shared plan be at least “publicly accessible,” namely, that the participants could discover the parts of the plan that pertain to them and to others with whom they are likely to interact if they wished to do so. In the interests of brevity, I have omitted this condition in the discussions that follow.
I, for example, share a plan to cook together because we designed the plan for us so that we may cook a meal, we each accept our parts and are committed not to undermine the other’s efforts.

On the planning theory, therefore, we can say that the constitutional law of the New York State is the shared plan that structures legal activity in New York because (1) the New York State Constitution was developed in 1938 by the New York State Constitutional Convention Committee so that a collection of individuals who meet certain qualifications can create and apply rules for the people of New York State, (2) those individuals intend to play the roles set out in the Constitution and the remaining parts of the State’s constitutional law, and (3) are committed not to interfering with others playing their respective roles. Those who accept New York State’s constitutional law (in the sense of (2) and (3)) are members of the New York State legal system and acts together with all others who accept the same rules.

Notice that sharing a plan, and hence acting together according to that plan, does not require that the participants care at all about the success of that plan, or even intend that their actions contribute to its success. Legal officials may be completely alienated from their roles; judges may apply the law simply in order to advance their careers, to avoid criminal sanctions or to pick up their pay checks. As long as the fundamental rules of the system were designed for individuals like them, the officials intend to do their part and not to interfere with other officials doing their parts, and they act on their intentions, we may say that they share a plan and act together in governing their community.

V. RESPONSES TO OBJECTIONS

56 By claiming that all legal systems are structured by a shared plan, I do not mean to suggest that all legal systems have been designed in advance. Historically, certain fundamental aspects of legal systems have arisen purely through custom. The model of plan-sharing I set out in the text above accounts for these cases by requiring only that the shared plan be designed “at least in part” with the group in mind. Groups may share plans, in other words, even though parts of their plans have not been planned for the group. A plan is shared if at least some part of the plan was designed for the group and group members see the non-planned parts as means to carry out the ends of the shared activity. Thus, the shared plan of a legal system may contain many customary parts, so long as it also contains non-customary parts and the officials see the customary parts as sub-plans of these non-customary parts.

57 In “Law, Plans and Practical Reason,” I argued that legal officials act together only if most intend to contribute to the creation and maintenance of a unified system of norms. Scott J. Shapiro, Law, Plans and Practical Reason, 8 Legal Theory 387, 419-21(2002). I now believe this condition is too strong. On the revised view set out in the text above, legal officials need have no intention to contribute to the existence of their legal system. In order to engage in the shared intentional activity of legal regulation, there must at least be a shared plan (which does not require intentions to contribute to the goals of the plan) and the members of the group must act on that plan.
In this part, I would like to show how the planning theory of law can help address the three objections to Hart’ doctrine we examined in Part Three. While the solutions I offer blunt the main force of the canvassed challenges, it will be quickly apparent that not every aspect of Hart’s doctrine of the rule of recognition, or his theory of legal obligation and authority, can be salvaged in the process. The objections show that Hart’s particular jurisprudential vision is flawed, but the responses offered suggest that his basic positivistic picture of law and its fundamental rules remain viable.

Shared Plans and the Content of Legal Systems

According to our interpretation of Hart, the rule of recognition is a unitary norm that imposes a duty on officials to apply certain rules that bear certain characteristics. It follows that, on such an account, the law of a particular system consists of all the norms that this rule obligates officials to apply. The first problem with this view, as we have seen, is its under- and over-inclusiveness: some rules which are part of the same legal system are not so considered by Hart’s account and, conversely, some of the rules which are considered part of the same legal system are not in fact part of it.

These problems would be alleviated, I would like to suggest, if we widened our lens so as to privilege not only the duties of courts but the powers of legislators as well. On this proposal, the rule of recognition in the United States should be identified with all of the constitutional provisions that allocate rule-creating powers and impose rule-applying duties. Roughly speaking, a rule is a law of the United States just in case it was created in accordance with, and its application regulated by, American constitutional law. Congressional statutes are thus part of the American legal system because they were created by Congress and the President in accordance with Articles I and II and federal and state officials are under a duty to apply these laws in their official capacities.

My suggestion that the rule of recognition be identified with the norm-creating and –applying provisions of a system’s constitutional order is not meant to exclude ordinary legislation from being part of the criteria of validity. In the United States, for example, statutory provisions such as the Judiciary Act of 1789 and the Administrative Procedure Act are bona fide elements of the rule of recognition. Even though they are not formally part of the United States Constitution and hence not entrenched

---

58 Clearly, the rule of recognition will no longer be an ultimate rule on this conception, although portion of it will be.
59 This characterization is not quite accurate because it does not cover binding customs, which are not created in accordance with power-conferring norms. In these cases, customary rules are part of a legal system just in case officials are under a duty to apply them.
from revision, they nevertheless confer powers to create rules and impose
duties to apply them and hence should be understood as partially
constituting the criteria of validity for the US system.

To see how this proposal solves the problem of under-inclusiveness,
let us return to the case where the New York State Legislature and
Governor each create a rule. According to the planning theory, two rules
are part of the same system just in case they are created in accordance
with, and their application regulated by, the system’s shared plan. Since
these two rules were each created in accordance with, and their application
regulated by, New York State constitutional law, it follows that there are
both part of New York State law, which is the correct result.

Another virtue of this proposal is that it does not transmute rules of
one system into the rules of another merely because one group is under a
duty to apply the other’s rules. For two enacted rules to be part of the
same system they must have been created according to the power-
conferring provisions of the same shared plan. Thus, even though the
shared plan that structures New York legal activity requires that the rules
of New Jersey be applied in certain instances, New Jersey officials do not
share this plan with New York officials and hence the New Jersey rules
have not been enacted pursuant to the same plan as the New York rules.

It will surely be objected that my account which identifies the rule of
recognition with the rule-creating and –applying portions of a system’s
shared plan marks no advance over Hart’s, for I have given no justification
for supposing that the shared plan that structures legal practice is one
shared plan, not many shared plans.

Two responses are in order. First, there is a very good reason to
suppose that the fundamental rules that set out the rights and
responsibilities of legal participants are all part of the same shared plan.
Because the shared plan is a group plan, it should be individuated
according to the group whose conduct it is supposed to guide. And since
the fundamental rules of a legal system are designed (at least in part) for
the group of officials of that system, it is natural to treat them as forming
one plan, not many. Thus, the rules that empower the Governor of New
York and the rules that empower the New York State Legislature are sub-
plans of the same plan because each was conceived as defining the role of
part of the group that creates and administers the laws of New York.

Secondly, and more importantly, although the account of a legal
system’s content I sketched requires that there be one shared plan, it is
possible to relax this requirement without changing the essentials of the
approach. Thus, we could say that a law is a member of a legal system
just in case there is a set of plans that a group share and stipulate that a
group shares a set of plans if and only if the plans of the set were
designed, at least in part, for them so that they may engage in a joint
activity, the members of the group accept the parts of the plans that apply
to them and are committed not to interfering with the parts that apply to
others.
On the planning theory, therefore, the content of a legal system does not depend on the unitary nature of its shared plan. Rather, it ultimately depends on the fact that someone conceived of officials as a group and developed a set of instructions to them so that they may collectively govern a community. The normative unity of law, we might say, depends on the social unity of officials. Whether we treat the instructions addressed to this group as forming one plan or many is, in the end, immaterial.

The Normativity of Law

As I have argued, Hart’s theory is unable to characterize accurately the content of a legal system because it focuses too narrowly on a small part of the constitutional structure. On Hart’s myopic view, it is immaterial that congressional legislation has been enacted in accordance with Articles I and II, insofar as these provisions are power-conferring, not duty-imposing. This, we have seen, is a mistake. Legislation counts as law of a particular system in part because it was created in accordance with the shared plan that structures the collective activity of legal regulation. Articles I and II are major parts of the American plan of governance and hence are essential for characterizing the content of American law.

Yet, one might object that our solution to the first objection precludes us from responding to the second objection. Since shared plans are social norms – they exist and are shared by a group just in case they have been designed with the group in mind and are accepted by each member of the group – they raise the same difficulty that the secondary rules raised for Hart, namely, the problem of normativity. How, for example, can anybody have legal authority to impose obligations simply because certain of their cronies authorize them to have such a power and members of the community acquiesce? Similarly, how can judges be under a legal obligation to apply certain rules just because other judges plan to do so as well? To generate normative relations of legal authority and obligation, the objection goes, a group needs more than social facts – it needs moral facts as well.

The proper response to the second objection, I believe, is to concede that the shared plans which constitute legal practice do not necessarily confer rights and obligations. What they do always succeed at doing, however, is to confer legal rights and legal obligations, which may (or may not) coincide with actual rights and obligations. And as long as one can show that shared plans are capable of generating legal rights and obligations, then the planning theory is able to account for the normativity of law.

In order to explain what I mean, let us begin with a basic question: when we attach the word “legal” to terms like “obligation,” “right,” “wrong,” “authority” and so on, what are we doing? One possible answer is that “legal” acts as an adjective modifying the noun phrase that follows.
A legal obligation is an obligation that is *legal*, namely, one that arises from the operations of legal institutions. On this interpretation, then, a legal obligation is an obligation that one has because of the law.

According to a second interpretation, the word “legal” acts as a qualifier, not a modifier. To say that one has a legal obligation, for example, is simply to assert that *from the legal point of view* one has an obligation. Statements of legal obligation, on this interpretation, are perspectival assertions. Regardless of whether one believes that the law has created actual obligations or has existing authority to do so, when one claims that another has a legal obligation, one is making an assertion from the point of view of the law. From the law’s perspective, it has the actual authority to impose actual obligations.

What, then, is the legal point of view? It is not necessarily the perspective of any particular legal official. No official may accept the law’s conception of itself. The legal point of view, rather, is the perspective of a certain normative theory. According to that theory, those who are authorized by the norms of legal institutions have moral legitimacy and, when they act in accordance with those norms, they generate a moral obligation to obey. The legal point of view of a certain system, in other words, is a theory that holds that the norms of that system are morally legitimate and obligating. Thus, communism is the point of view of communist legal systems, individualism the point of view of laissez-faire capitalist systems, democratic theory the point of view of democratic systems and so on.

The normative theory that represents a system’s point of view may, of course, be false from a moral perspective. That is, the legal point of view may not coincide with the true moral point of view. Those authorized by legal institutions to act may be morally illegitimate and their actions may generate no moral obligations to obey. The point of view of a particular legal system may be like the phlogiston theory of combustion: a scientific theory that aimed to be true but missed the mark. In short, the legal point of view always *purports* to represent truly the moral point of view, even when it fails to do so.

I would like to suggest that when we say that the law necessarily has the power to confer legal rights and impose legal obligations, we are using the word “legal” in the second, qualifying sense. We are *distancing* ourselves from our normative assertions, claiming only that from the legal point of view the law’s activities are reason-giving. On this second interpretation, it is easy to see how even morally illegitimate shared plans can confer legal rights and impose legal obligations. For to ascribe legal authority to a body in a particular legal system is to assert that, from the point of view of that legal system, the body in question is morally legitimate.

(1) X has legal authority over Y in system S $\iff$ From the point of view of S, X has moral authority over Y.
The point of view of that legal system, in turn, will ascribe moral legitimacy to a body just in case its norms confer legal power on that body. Since on the planning theory the legal norms that confer legal authority are subplans of the system’s shared plan, the legal point of view will ascribe moral legitimacy to a body when its shared plan authorizes that body to so act.

(2) From the point of view of S, X has moral authority over Y
\[ \iff \text{The shared plan of S authorizes X to plan for Y.} \]

It follows from (1) and (2) that a body will have legal authority in a particular legal system just in case the system’s shared plan authorizes that body to so act.\(^61\)

(3) X has legal authority over Y in system S \[ \iff \text{The shared plan of S authorizes X to plan for Y.} \]

Contrary to the second objection, then, accounting for the normativity of law does not require showing that the secondary legal rules are always capable of creating rights and obligations. One must only demonstrate how the existence of the secondary rules necessarily ground normative judgments made from the legal point of view. As we have seen, the shared plan of a legal system renders true certain perspectival judgments even if the shared plan happens to be morally illegitimate. For a body has legal authority in a system, and thus the ability to impose legal obligations, just in case the shared plan authorizes it and a shared plan authorizes such a body just in case certain social facts obtain.

\textit{Law and Disagreement}

According to the third objection, Hart’s doctrine of the rule of recognition must be flawed because it cannot account for pervasive, well-known but sincere disagreements about the ultimate criteria of legal validity. As Dworkin pointed out, widespread disagreements about the content of the rule of recognition are inconsistent with the consensus which supposedly generates its content. Thus, if legal participants are neither hopelessly confused about legal practice nor opportunistic liars, the criteria of legal validity cannot be determined by judicial agreement about those very criteria.

One reaction to Dworkin’s objection (Dworkin’s reaction, in fact) is to deny that the ultimate criteria of legality can ever be determined by

\(^{61}\) X will have legal authority over Y in S only when S is generally efficacious. Hence, it will not be sufficient for the shared plan to authorize someone to plan in order for that person to have legal authority. It must be the plan of a generally efficacious planning system.
existing official consensus. This response, I think, would be too hasty. For one extremely appealing aspect of Hart’s theory is how it discounts the importance of what Woodrow Wilson once called the “literary Constitution” in favor of the “Constitution in operation.” By privileging current social practice, Hart’s theory is able to account for the legality of actions that would otherwise be very difficult to justify. For example, the Supreme Court has held that the Due Process clause of the Fifth and Fourteenth Amendments should be interpreted substantively, as well as procedurally. Under so-called “substantive” due process analysis, government must not only provide fair procedures for the adjudication of legal claims, but ensure that the individuals are afforded certain basic rights as well. To say the very least, this interpretation is quite strained. As John Hart Ely once quipped, “there is simply no avoiding the fact that the word that follows ‘due’ is ‘process’ … ‘Substantive due process’ is a contradiction in terms — sort of like ‘green pastel redness’.” Despite its apparent absurdity, this interpretation of the text is now legally correct. It is so because most everyone currently accepts that the Constitution confers a right of substantive due process on individuals. End of story.

Present consensus, therefore, should be seen as a sufficient condition for determining the ultimate criteria of legal validity. What Dworkin’s critique of Hart shows, I believe, is that it cannot be a necessary condition. In some instances, there may be a fact of the matter as to whether a certain test is legally proper despite the lack of agreement on such a question.

It is important to see that this acknowledgment is consistent with the core positivistic claim that the existence and content of the rule of recognition are determined by social facts alone. Specific agreement on the criteria of validity counts as a social fact for these purposes, but such consensus is only one kind of social fact. It is possible that the rule of recognition will be determined by social facts other than agreement on its existence or content. It is this possibility I would now like to pursue.

Let us devote our attentions to American-style legal systems, namely, ones which have developed through a self-conscious process of constitutional design. In any such regime, there will usually be an existing agreement on at least three constitutional matters: 1) its basic institutional arrangements; 2) those empowered to affect its structure (which we might

---

62 Another reaction is to claim that there can be multiple inconsistent rules of recognition in a particular system, each one determined by the sub-group which accepts it. For this possibility, see Matthew Adler, Popular Constitutionalism and the Rule of Recognition: Whose Practices Ground U.S. Law?, 100 Northwestern Law Review 719 (2006).

63 WOODROW WILSON, CONGRESSIONAL GOVERNMENT 30 (Johns Hopkins Univ. Press 1981) (1885). As American constitutional theorists would now put the point, Hart discounts the big ‘C’ Constitution (the document) in favor of the small ‘c’ constitution (i.e., the practice of constitutional law). On this distinction, see e.g. David A. Strauss, The Irrelevance of Constitutional Amendments, 114 Harv. L. Rev. 1457, 1459-60 (2001).

64 JOHN HART ELY, DEMOCRACY AND DISTRUST 18 (1980).
call its “constitutional designers”); and 3) its authoritative texts. Since present consensus is a sufficient condition for determining the existence and content of a shared plan, these agreements partially specify the shared plan of that legal system. In order to figure out the remainder of the system’s shared plan, an interpreter must ascertain the proper way to interpret the authoritative texts which set it out.

Of course, a consensus might exist in this system about which interpretive methodology ought to be used, in which case the agree-upon methodology would indeed be legally authoritative for such a regime. Yet, what does the interpreter do when she works within a system, like the one in the United States, in which there is no official accord on interpretive methodology?

The proposal is that the proper way to interpret these texts can be derived by focusing on the reasons that the system’s constitutional designers had for adopting its basic institutional arrangements and using these reason to figure out which interpretive methodology would best harmonize with these reasons. In particular, the interpretive methodology that best furthers the designers’ shared goals, values and judgments of trustworthiness is the proper one for interpreting the authoritative texts and hence for revealing the content of the system’s shared plan.

An example may help motivate this procedure. Consider a regime in which the constitutional designers hold a very distrustful view of the competence and character of officials. As a result, they aim to create a certain legal framework for coping with such problems: they intend to diffuse authority through the system, forbid executive and judicial officers from legislating, set up lengthy waiting times before legislation can be passed, enforce sanctions for abuse of discretion, …, etc. They also draft a constitution that sets out these rights and duties in very clear and precise language. Suppose further that after ratification, there is a general consensus among officials about the basic institutional arrangements of the regime. That is, everyone accepts that executive and judicial officers are forbidden from legislating, there are lengthy waiting times before legislation may pass, sanctions should imposed for abuses of discretion, certain individuals have the authority to alter these arrangements and so on.

In contrast to the jaundiced views of the designers, however, the officials who must interpret the constitution think of themselves as eminently trustworthy. They believe that the constraints placed upon them by the constitutional designers are unnecessary and impede their valuable work. Hence, when they interpret the texts that set out the rules of the system, they use their liberal views about their own trustworthiness and

---

65 In the American system, for example, the constitutional designers ordinarily include Congress, the President, state legislatures, constitutional conventions, and federal courts. While there is not universal agreement about the entire roster of constitutional planners (e.g., are We the People designers?), there is I believe a list that one could draw up that would command sufficient consensus among American jurists.
assume large degrees of discretion in interpretation: they read grants of power broadly, interpret constraints narrowly, ignore legislative texts when it gives a result with which they mildly disagree, refuse to defer to the interpretation of regulations by the appropriate administrative agencies, etc.

The obvious difficulty with this mode of proceeding is that the very point of having designers design the constitutional order is undone by the actions of the interpreters. The shared plans that set out the distribution of rights and responsibilities are supposed to resolve second-order uncertainty in general, and questions of trust in particular. However, if the interpreters are authorized to use their own judgments of trustworthiness in order to determine interpretive method, and to use that method to interpret legal texts, then they defeat this aim. Whenever the current designers want to constrain discretion, they can widen discretion; when the current designers want to widen discretion, they can constrain discretion. It is the interpreters’ views that ultimately determine the system’s economy of trust, not the plan or the current designers. Here, Bishop Hoadly’s famous dictum is apt: “Nay, whoever hath an absolute authority to interpret any written or spoken laws, it is he who is truly the Law-giver to all intents and purposes, and not the person who first wrote or spoke them.”

Thus, if the shared plan of a legal system is to resolve political issues relating to goals, competence and character, its content cannot depend in any way on the goals that the system morally ought to pursue or the competence and character that legal officials truly possess. And since the content of a shared plan depends on the correct way to interpret the texts that set it ought, the proper way to determine interpretive methodology cannot depend in any way on which goals are morally best or the actual trustworthiness of officials.

As the above example suggests, determining proper interpretive methodology cannot be a “protestant” affair. The fatal defect of this approach is that it is self-defeating, namely, it makes no sense from an organizational perspective to empower the current designers of legal systems to control the system’s authority structure and the content of its legal texts, but not its interpretive method. For any attempt to resolve second-order uncertainty through the process of institutional design and legislative drafting would be defeated at the stage of interpretation. Those questions of morality and trust that were settled at ratification would suddenly be reopened during implementation. And, by allowing interpreters’ views on morality and trust to determine interpretive method, and how much interpretive discretion they should be allowed, the

---

67 On the “protestant” attitude towards the law, see Dworkin, supra note 49 at 413.
implementers are able to substitute their attitudes for those of the designers.

It is easy to see, I think, that the same logic which excludes moral and psychological truths as the determinants of proper interpretive methodology necessitates that the judgments of those who are deemed the appropriate designers for constitutional matters should control. For if their assessments about, say, trustworthiness do not control the proper way to interpret texts which set out shared plans, they cannot control the distribution of political power, and hence resolve second-order uncertainty, which is the very point of having them fashion shared plans in the first place. If the constitutional designers are distrustful of officials, interpreters must take these judgments as given for the purposes of legal interpretation lest they arrogate to themselves too much power from the legal point of view. Conversely, trusting attitudes should lead greater interpretive discretion; otherwise, legal participants will preclude themselves from pursuing the objectives that they were entrusted to serve.

Determining Interpretive Methodology: The Case of Originalism

According to the planning theory, the proper interpretive methodology for a legal system that has been (1) self-consciously planned (2) by a group of agreed-upon constitutional designers is the procedure which best furthers the goals and values that the system has been designed to serve in light of the attitudes of trust which motivated the distribution of political power. Needless to say, the relationship between the shared ideology of constitutional designers regarding goals, competence and character and proper interpretive methodology is highly complex. Setting out the many complex links that exist between them is clearly beyond the scope of this paper. But I would be remiss if I did not give the reader an approximate sense of how the procedure works in practice.

Roughly speaking, the planning theory requires that interpretive discretion track systemic judgments of trustworthiness: an interpretive methodology that requires for its effective implementation a high degree of competence or moral character will be inappropriate for systems designed in accordance with distrustful views of human nature; instead, hermeneutic procedures that are easier to apply and less subject to abuse – perhaps ones that defers to plain meaning, instead of purpose – would be more fitting.

Indeed, popular arguments for theories of constitutional interpretation that privilege framers’ intentions can be seen on the model that I have presented, namely, that the American legal system, in one way or another, is distrustful of individuals and that the best way to deal with this distrust is to confine legal interpretation to original understanding. For example, in A Matter of Interpretation, Justice Antonin Scalia argues that constitutional provisions ought to be interpreted in accordance with

---

68 I explore these relationships in detail in SCOTT J. SHAPIRO, LEGALITY (forthcoming).
the original meaning of the text, rather than with an “evolving sense of decency.” A judge living today should not, for example, interpret the Eighth Amendment according to the meaning that she assigns to the term “cruel” if that meaning diverges from the late 18th Century understanding, for such evolutionary methodologies flout the fundamental function of constitutions.

“It certainly cannot be said that a constitution naturally suggest changeability; to the contrary, its whole purpose is to prevent change — to embed certain rights in such a manner that future generations cannot readily take away. A society that adopts a bill of rights is skeptical that ‘evolving standards of decency’ always ‘mark progress’ and that societies always ‘mature,’ as opposed to rot.”69

On Scalia’s view, then, the purpose of constitutions is to prevent untrustworthy future generations from rescinding rights that the present generation has deemed proper. But, he argues, granting judges the power to interpret the constitutional text in accordance with changing conceptions of morality would effectively permit future generations to change the constitution and thereby defeat its raison d’être. Scalia, therefore, argues for originalism by noting the distrustful nature of the American constitutional order and then claiming that living constitutionalism is inconsistent with this economy of trust.

Likewise, opposition to such views can be understood on the planning theory, for a standard rejoinder to the originalist claims of distrust is to argue that the American constitutional order is not nearly as wary of courts as these originalists suggest. Opponents point to the fact that the Constitution often eschews concrete and particular language in favor of setting out broad statements of moral principle, such as when it prohibits “cruel” punishment, mandates “due” process and guarantees “equal” protection of the law. This would seem to indicate that the framers trusted future generations to use their moral judgment in determining which state action is acceptable. As Dworkin has argued: “Enlightenment statesmen were very unlikely to think that their own views represented the last word in moral progress. If they were really worried that future generations would protect rights less vigorously than they themselves did, they would have made plain that they intended to create a dated provision.”70 Thus, some form of living constitutionalism would best harmonize with the distribution of trust and distrust manifested in the American constitutional order.

As this brief discussion indicates, although the planning theory requires deference to planner attitudes about goals, values and trust, it is not a version of originalism. First, originalism is an interpretive

70 Id. at 124.
methodology, whereas the planning theory sets out a decision-procedure for adjudicating between interpretive methodologies. The planning theory may, of course, recommend originalism in certain circumstances, namely, in situations where the designers’ attitudes of trust demand this. But, as we just have seen, it is entirely possible that attending to their trust attitudes requires that the original understanding of certain textual provisions be ignored.

Second, originalism focuses on original intent, that is, on the attitudes of those who framed particular texts. The planning theory, on the other hand, does not privilege the views of the system’s original constitutional designers. Because legal systems always contain mechanisms for revision, the constitutional designers change as the structure of the system changes. The designers of the present American system not only include the framers and ratifiers of the Constitution of 1787, but the numerous agents over the past two hundred years that have changed the complexion of the system. Moreover, the constitutional designers who are relevant for determining interpretive methodology are those singled out by the present consensus in the legal community. This present consensus determines which past consensus to heed. The idea, once again, is that it is irrational for a group to treat a set of agents as designers whose role is to resolve second-order uncertainty and at the same time not privilege their attitudes about appropriate goals, values and trust when trying to figure out how to interpret their instructions. Protestant practices, I have argued, are self-defeating and therefore cannot represent proper legal reasoning.

Social Facts Without Total Consensus

The advantages of the planning theory, I believe, are considerable. Chief among them is that, insofar as official consensus is not necessary for the determination of interpretive methodology, the planning theory is able to account for the possibility of disagreements about the ultimate criteria of validity. Participants in a practice can disagree over proper interpretive methodology because they disagree about the demands imposed by particular methodologies, the goals and values of the system, its economy of trust or which methodology best harmonizes with such ideologies.

In order to secure this result, as well as to respond to the other objections, we have seen that the planning theory departs from Hart’s doctrine in several important respects: it treats all of the norm-creating and –applying provisions of a system’s constitution, instead of merely a portion thereof, as its rule of recognition; denies that the secondary rules always confer rights and impose duties (as opposed to legal rights and duties); and deems present official consensus as merely sufficient, but not necessary, for the determination of the criteria of legality.

Nevertheless, I think that the planning theory is at least Hartian in spirit, if not in letter. First, like Hart’s theory, the planning theory does
not require that the fundamental rules of a legal system be morally desirable. The shared understandings of a legal community and the system’s animating ideology may either be ethically odious, scientifically backwards or both. Nevertheless, these considerations are taken as settled and are thus used to determine the ultimate criteria of legal validity.

Second, like Hart’s theory, the planning theory ultimately grounds the secondary rules in facts about the behavior and attitudes of groups. After all, that a group of constitutional designers shared a certain ideology regarding goals, values and/or trust is a social fact. Similarly, that a legal community presently shares an understanding about the identity of those designers, and the basic structure and texts they have created, is also a social fact. The shared plan of any legal system, then, is a social rule because its existence and content is determined by social facts alone.

Finally, both Hart’s rule of recognition and the planning theory’s shared plan play the same role, namely, the resolution of normative uncertainty. One could argue, in fact, that this concurrence on function is the most important one, insofar as any account that shares this equivalence is required to share the others properties as well. As I claimed in Part Four, doubts and disagreements concerning second-order questions of political morality are bound to be as socially confounding in communities governed by law as first-order ones that concern private obligations. A rule of recognition that exists simply in virtue of its moral desirability, however, cannot and will not resolve such disputes. For those who have doubts or disagree about who has legitimate authority would first have to know, or agree about, the moral facts and which marks of authority those moral facts pick out. By hypothesis, these parties neither know nor agree about these very issues.

By contrast, a rule whose existence and content was determined by social facts alone could resolve such doubts and disagreements. One would not have to know whether one was truly entitled to rule; one would simply have to know who was designated by the shared understandings or practice of the relevant legal participants and work from there.

CONCLUSION:
THE EXISTENCE OF THE RULE OF RECOGNITION

So, does the rule of recognition exist? Well, it all depends on what the rule of recognition is. If we take the rule of recognition in a very minimal manner – as the test of legal validity for a particular legal system – then everyone agrees that such a rule exist. Even “Law as Integrity” is a rule of recognition in this anodyne sense.

On the other hand, if we construe the rule of recognition as Hart did – as a duty-imposing convention among officials – then I think we must conclude that the rule of recognition does not exist. For as we have seen, such a rule cannot accurately characterize the content of a legal system;
impose duties or confer powers; or exist in the face of disagreement about its content.

Finally, if we take the rule of recognition of a legal system to be constituted by the norm-creating and applying provisions of its shared plan, then I believe that it does exist. Like Hart’s rule of recognition, this norm is always at least partially constituted by official convergence on a standard of conduct. But unlike Hart’s account, total convergence is not necessary. As long as there is present agreement among officials on the basic structure of the constitutional order, the constitutional designers and the authoritative texts and past consensus among the constitutional designers about the goals and values the institution is to serve and the degree of trust that is warranted to show to members of the community, then the raw materials are available from which proper interpretive methodology may be divined and, in turn, the remainder of the rule of recognition may be ascertained.

It no doubt follows from this account that there must be sufficient consensus about the content of the legal system in question in order for there to be a proper interpretive methodology to find. At the very least, there must be ample shared understandings about who the constitutional designers of the system, and the basic institutional structure and authoritative texts they have created. These happy convergences provide the pre-interpretive materials that form the heart of the system’s economy of trust and from which the determination of interpretive methodology must proceed. Without them, the procedure cannot get off of the ground.71

Lack of consensus, of course, does not preclude actors from arguing about appropriate interpretive methodology. As long as disputants think that there is such convergence, or at least act as though they do, each side can fashion, against this assumed common ground, coherent arguments for originalism, interpretivism, pragmatism or whichever –ism they support. The absence of presupposed consensus merely precludes either side from being correct. Their hermeneutical disputation may be filled with sound and fury, but from the legal point of view they signify nothing. In these cases at least, I believe that Hart’s description of fundamental constitutional controversies is correct. “Here, all that succeeds is success.”72 A misguided legal argument, or covert political argument, may catch on and be taken as true by the legal community. Should this happens, the embraced political position will be transformed into a true legal conclusion and the plan that they all share will shift accordingly.

71 I leave it as an open question whether there is another procedure that will determine interpretive methodology in the absence of the convergences mentioned in the text.
72 CL 153.