



10-20-2014

Legal Formalism, Procedural Principles, and Judicial Constraint in American Adjudication

Christopher J. Peters

University of Baltimore School of Law, cpeters@ubalt.edu

Follow this and additional works at: http://scholarworks.law.ubalt.edu/all_fac



Part of the [Law Commons](#)

Recommended Citation

Christopher J. Peters, *Legal Formalism, Procedural Principles, and Judicial Constraint in American Adjudication*, (2014).

Available at: http://scholarworks.law.ubalt.edu/all_fac/880

This Article is brought to you for free and open access by the Faculty Scholarship at ScholarWorks@University of Baltimore School of Law. It has been accepted for inclusion in All Faculty Scholarship by an authorized administrator of ScholarWorks@University of Baltimore School of Law. For more information, please contact snolan@ubalt.edu.

Legal Formalism, Procedural Principles, and Judicial Constraint in American Adjudication

Christopher J. Peters

University of Baltimore School of Law, Baltimore, MD, USA (cpeters@ubalt.edu)

Abstract American proponents of legal formalism, such as Supreme Court Justice Antonin Scalia, worry (quite reasonably) that unfettered judicial discretion poses a threat to democratic legitimacy, and they offer formalism – the mechanical implementation of determinate legal rules – as a solution to this threat. I argue here, however, that formalist interpretive techniques are neither sufficient nor necessary to impose meaningful constraint on judges. Both the text and the “original meaning” of legal rules are endemically underdeterminate, leaving much room for judicial discretion in the decision of cases. But meaningful judicial constraint can and does flow from other sources in American adjudication. Judges are constrained by the dispute-resolving posture of their task, which requires that they be impartial as between the litigants and responsive to the litigants’ participatory efforts. And they are constrained by the need to be faithful to the substantive principles that justify legal rules, even when those rules themselves are indeterminate. Judicial constraint in the American system thus stems not primarily from formalist interpretative methods, but rather from largely unwritten procedural principles of judicial impartiality, responsiveness, and faithfulness.

Now the main danger in judicial interpretation of the Constitution – or, for that matter, in judicial interpretation of any law – is that judges will mistake their own predilections for the law. ... [¶] Originalism does not aggravate the principal weakness of the system, for it establishes a historical criterion that is conceptually quite separate from the preferences of the judge himself.¹

Hon. Antonin Scalia
Associate Justice, United States Supreme Court

1 “The Main Danger”: Legal Formalism and Judicial Constraint

The quotation above from Justice Scalia, perhaps the most prominent proponent of legal formalism in the United States, articulates an intuitive connection between formalism and judicial constraint: to the extent judges are limited to the rote application of existing rules, they are prevented from deciding cases according to “their own predilections” or “preferences.” Justice Scalia’s approach implies that formalist principles are both sufficient and necessary to limit judicial discretion. In this chapter, I argue that they are neither.

For the purposes of this chapter, we can understand *legal formalism* as *the idea that judges and other decisionmakers should decide particular cases, to the extent possible, by the mechanical application of existing legal rules*. Those existing rules might come from constitutional provisions, from statutes, from treaties, from administrative regulations, or from the decisions of prior courts. To the extent a judge’s decision of a case is dictated solely by the content of an existing rule, that decision is not determined by anything else, including the values or beliefs or preferences of the judge herself. There is, then, a quite literal relationship of semantic or logical entailment between pure legal formalism and judicial constraint. A purely formalist decision is an entirely constrained one – a decision constrained completely by the content of the existing rule being applied.

Why might constraining judges and other point-of-application decisionmakers be a good idea? Judges constrained by formalism will not always produce the best possible decisions; sometimes the rule being applied will be a bad rule, and sometimes even the application of generally good rules will produce bad results, as Aristotle understood.² There may nonetheless be strong reasons of rule-

¹ Scalia (1989a, 863-64).

² In the *Nicomachean Ethics*, Aristotle (1941, 1020) wrote:

[A]ll law is universal but about some things it is not possible to make a universal statement which shall be correct. In those cases, then, in which it is necessary to speak universally, but not possible to do so correctly, the law takes the usual case, though it is not ignorant of the possibility of error. And it is none the less correct; for the error is not in the law nor in the legislator but in the nature of the thing, since the matter of practical affairs is of this kind from the start.

consequentialism to require judges to apply existing rules even where the judge believes the result would be bad.³ No doubt there also are “rule of law” reasons such as predictability and consistency.⁴ In this paper, however, I want to focus on the sorts of reasons to which Justice Scalia primarily appeals in his defense of legal formalism: reasons of *democratic legitimacy*.

As Cass Sunstein (1997, 530) has written:

Justice Scalia is a democrat in the sense that much of his jurisprudence is designed to ensure that judgments are made by those with a superior democratic pedigree. Above all, he seeks to develop rules of interpretation that will limit the policymaking authority and decisional discretion of the judiciary, the least accountable branch of government.

The American judiciary – in particular, the federal judiciary – is “the least accountable branch of government” in Justice Scalia’s view because its members serve during “good behavior” (that is, in most cases, for life or until they choose to retire),⁵ thus insulating their decisions from electoral accountability. For Justice Scalia, it is significant that the American judiciary appears to be saliently less accountable, politically speaking, than most of the institutions whose rules it is charged with interpreting and applying: the legislature, state or federal (whose members are directly elected by the voters); the administrative bodies whose chief policymakers typically are appointed and confirmed by elected officials; and the constitutional framers, whose efforts were channeled through an extraordinarily deliberative and participatory political process.⁶

If we take democratic accountability as a standard of political legitimacy, as Justice Scalia does, then the judicial deficit in accountability as compared to these other decisionmakers renders judicial decisions less legitimate than political decisions, all else being equal. To be precise, it renders decisions based on judicial *discretion* – departing from the law or creating new law as opposed to simply ap-

³ For an argument to this effect, focused on judicial application of rules gleaned from precedent, see Hellman (2014).

⁴ See, for example, Scalia (1989b, 1179).

⁵ See U.S. Const. art. III, § 1: “The Judges, both of the supreme and inferior Courts, shall hold their Offices during good Behaviour, and shall, at stated Times, receive for their Services a Compensation, which shall not be diminished during their Continuance in Office.”

⁶ The obvious apparent exception is the prior judges whose decisions often are applied as precedents by current judges. Where these precedents interpret the work of institutions that are politically accountable to current majorities – statutes enacted by legislatures and regulations adopted by administrative agencies – their continued existence might be seen as implicit acquiescence in their substance by these accountable institutions. (The legislature could simply overturn incorrect interpretations of its statutes by amending the statute in question.) On this theory, respect for prior judicial decisions might be understood as a form of subservience to democratically more-accountable institutions. Matters are more obscure where the judicial precedent interprets a constitutional provision and thus is very difficult to correct by means of constitutional amendment. Many formalists thus distrust the presumptive American practice of adhering to constitutional precedent (see Peters (2014, 189-98), although Justice Scalia himself professes to accept the practice (see Scalia (1989a, 861): “[A]lmost every originalist would adulterate [originalism] with the doctrine of *stare decisis*”).

plying existing law – less legitimate than the discretionary decisions of politically accountable actors. To hold judges strictly to an application of the rules created by these more-accountable decisionmakers thus is, for Justice Scalia, to allocate lawmaking authority to democratically superior institutions (legislatures, constitutional framers) as opposed to democratically inferior ones (courts). Legal formalism promotes democratic legitimacy.

The view that Justice Scalia represents therefore implies judicial constraint from legal formalism and democratic legitimacy from judicial constraint. For purposes of my arguments here, I will grant the premise that some meaningful degree of constraint upon judicial discretion is necessary for the proper functioning of democratic governance. But I will contend that the most important sources of this constraint lie in principles other than those urged by legal formalists.

2 “A ... Criterion ... Quite Separate from the Preferences of the Judge,” Part I: The Endemic Indeterminacy of Text

In order for legal formalism to work as advertised – to constrain judges – it must be capable of doing so. Judges must actually be able to decide most or all cases primarily or exclusively by the mechanical application of existing legal rules. And whether or not judges are deciding cases in this way must be transparent to others – to the litigants, to superior judges (if any) in the hierarchy, to policymakers, to lawyers, to legal academics, perhaps to the media and to the public – if the constraint imposed by legal rules is to be real rather than merely professed. If it is not clear in most cases whether judges are in fact simply applying existing rules, then judges often will be able to avoid detection in not doing so, thus substantially reducing their incentive to simply apply existing rules.

The constraining function of formalism therefore depends on the existence of a system of legal rules that is *determinate* – capable of conclusively resolving all legal issues in a particular case – and whose determinacy is transparent. Most legal rules are communicated in textual form, so it makes sense to ask whether the text of legal rules is capable, by itself, of conferring this sort of systemic determinacy. For a familiar set of reasons, the answer is no.

We can illustrate why using an example that will be familiar to students of Anglo-American jurisprudence, arising as it does from a well-known mid-twentieth-century debate between the English legal positivist H.L.A. Hart and the American “Legal Process” theorist Lon Fuller (see Hart 1958, 607; Fuller 1958, 662-63). Suppose a city ordinance prohibits “vehicles” in the public park. A group of war veterans wants to erect a monument in the park featuring a now-inoperative truck once used in combat. Would this violate the ordinance?

Note, first of all, that the applicable text of the ordinance – banning “vehicles” from the park – will not, standing alone, answer this question. As Hart saw it, any word has both “a core of settled meaning” and “a penumbra of debatable cases in

which [the word is] neither obviously applicable nor obviously ruled out” (Hart 1958, 607). For the latter category, Hart had in mind examples like whether “bicycles, roller skates, or toy automobiles” qualified as “vehicles” under the ordinance (id.). Fuller, for his part, was skeptical of the existence even of “a core of settled meaning”; he used the war-monument example to suggest that even something we normally would call a “vehicle” without a second thought can, in some instances, fall within Hart’s “penumbra” of ambiguity (Fuller 1958, 663). Both theorists understood, however, that there will be many cases in which the proper application of a rule’s text to the facts of a particular case will be uncertain. As Hart put it, “[t]he toy automobile” (or for that matter the inoperative combat truck) “cannot speak up and say, ‘I am a vehicle for the purpose of this legal rule’” (Hart 1958, 607).

The hypothetical “no vehicles in the park” ordinance thus is an example of an underdeterminate textual expression of a legal rule. Based upon the text alone, a judge faced with a case involving the war-monument combat truck could reasonably reach either alternative conclusion about the meaning of the ordinance: that the truck *is* a “vehicle” and thus is barred from the park, or that the truck is *not* a “vehicle” and thus is permitted. Neither interpretation would clearly be an unfaithful application of the text of the rule. And thus the judge is not constrained by the text alone to reach one interpretation rather than the other.

Nor is the “vehicles” example anomalous as a representation of textual indeterminacy in American law (or, I suspect, in the law of any reasonably complex legal system). In the context that most concerns Justice Scalia – American constitutional law – vague or ambiguous text is more the rule (as it were) than the exception. The “penumbras” of terms like “the equal protection of the laws” in the Fourteenth Amendment, “liberty” and “due process of law” in the Fifth and Fourteenth Amendments, and “the freedom of speech” in the First Amendment are considerably larger than that of the word “vehicles” in Hart’s hypothetical ordinance.

Even relatively determinate legal texts have instances of indeterminacy. Justice Scalia himself (Scalia 1997, 23-24) discusses an example of this (although he denies that it is an example): a federal statute mandating increased jail time for a defendant who “uses ... a firearm” “during and in relation to ... [a] drug trafficking crime.”⁷ Suppose a defendant offers to trade an unloaded gun for drugs; has he “use[d] ... a firearm” in the sense meant by the statute? In *Smith v. United States*, the Supreme Court answered yes; Justice Scalia dissented on the ground that the decision was inconsistent with the “ordinary meaning” of the phrase “uses a firearm.”⁸ The Justice won two other votes with his dissent; a six-Justice majority disagreed with his interpretation of the statute’s text. That the text was susceptible to at least two reasonable, and mutually exclusive, interpretations seems obvious from the nonunanimous vote.

⁷ 18 U.S.C. § 924(c)(1).

⁸ 508 U.S. 223, 241, 242 (1993) (Scalia, J., dissenting).

Add to this the fact that legal rules rarely stand alone in any legal system. Often they interact with other legal rules that apply in particular cases. For example, in a recent lawsuit challenging the validity of a federal law refusing recognition to same-sex marriages, the Supreme Court had to interpret not only the statute being challenged, but also the Equal Protection Clause of the federal Constitution, the Due Process Clause of that Constitution, and many prior judicial decisions applying those clauses.⁹ And it is virtually always true that legal rules are situated within a larger matrix of rules that, while not directly applicable to the case at hand, nonetheless may influence which interpretations of the applicable rule are most reasonable. Calling an inoperative war-memorial truck a prohibited “vehicle” will seem more or less reasonable to the extent it is consistent or inconsistent with, for example, the use of the word “vehicle” in other ordinances, or with a separate ordinance promoting “natural” uses of the park. A judge applying the indeterminate text of any given legal rule therefore often must also apply the indeterminate text of other rules and determine how the rules interact with each other.

So, in many cases within a complex legal system like that in the United States, the text of legal rules standing alone will be indeterminate, or at least will not be transparently determinate. Judges will be able to reach more than one reasonable result that is consistent with the text. This is not to say that text does not constrain at all in such cases; often there will be many applications of the text clearly within (or outside) its “core of settled meaning.” An operative combat truck driven through the park clearly would violate the “no vehicles” ordinance; a toddler’s small wooden pull-toy clearly would not. Nor is it to deny that some textual expressions of rules will be more determinate than others, or even that some will be determinate in every or nearly every case. (Consider a law setting the speed limit at 65 mph, or the provision of the U.S. Constitution composing the Senate of “two Senators from each State.”) The point is only that there will be a great many cases in which the text does not completely constrain judges. In many cases, the formalist search for a criterion of decision “quite separate from the preferences of the judge himself” (in Scalia’s words) will have to extend beyond the text.

3 “A ... Criterion ... Quite Separate from the Preferences of the Judge,” Part II: The Endemic Indeterminacy of Originalism

If text alone rarely can underwrite formalism, then what? Among American legal formalists, the answer is almost always some form of *originalism*. Originalists seek to supplement indeterminate legal texts with empirical facts about the process of the text’s adoption: the “intent” of the authors or ratifiers of the text, or – in the currently dominant form of originalism – the “meaning” that would have been attributed to the text by a reasonable member of the public at the time of its adoption

⁹ The decision in question is *United States v. Windsor*, 133 S. Ct. 2675 (2013).

as law. Justice Scalia, a leading originalist, was an early proponent of the latter approach (Scalia 1989a; 1997). This “original meaning” or “original public meaning” originalism superseded its “original intent” progenitor in part because the former was exposed as highly indeterminate,¹⁰ and so I will focus on the supposedly more-determinate recent version here.

For formalists, the point of a search for “original meaning” is to identify some factual determinant of legal meaning that is independent of a judge’s own values or desires. If, as a matter of historical linguistics, there is an “original public meaning” of, say, the Free Speech Clause of the Constitution’s First Amendment – “Congress shall make no law ... abridging the freedom of speech” – then a judge can apply that original meaning without relying on her own preferences. And if the original public meaning is relatively accessible and transparent, then the judge can be *seen* to be applying that meaning without relying on her preferences – or not to be doing so. The original public meaning thus promises to constrain judges.

The problems with this aspiration, however, are threefold. Often an original public meaning will not exist, at least not at the level required to decide a concrete case. Sometimes, to the contrary, there will be multiple inconsistent original public meanings. And locating a single original meaning, assuming it exists, frequently will be difficult or impossible, especially for judges, who are not in fact linguistic historians.

Often an original public meaning will not exist in the required sense because, at the time the text was adopted as law, nobody considered what that text might mean in a given circumstance. Sometimes nobody at the time of adoption *could* have considered the question. For example, when the Free Speech Clause became part of the Constitution in 1791, no one could have thought about its potential application to campaign-finance laws in the late-twentieth and early twenty-first centuries. It would have been impossible to anticipate what American political campaigns would look like two centuries later – the existence of television and the Internet, the dominance of two entrenched political parties, the system of primary elections, the importance of independent “political action committees,” the rise of for-profit corporations. To ask a reasonable and well-informed citizen in 1791 whether restrictions on so-called “issue advertisements” funded by corporations or unions would “abridge[e] the freedom of speech” would be to ask a nonsensical question. The words of the Free Speech Clause simply *have* no “original meaning” with respect to that question.

On the other hand, sometimes the evidence will suggest contradictory original meanings. Within a few years after ratification of the Bill of Rights, the Federalist Congress enacted the Sedition Act, which made it a federal crime to “write, print, utter, or publish ... any false, scandalous and malicious writing or writings against

¹⁰ For an originalist’s account of the general rejection, among originalists, of “original intent” originalism in favor of “original public meaning” originalism, see Solum (2011, 6-17). For influential critiques of “original intent” originalism as indeterminate, see Brest (1980); Dworkin (1985, 34-57).

the government of the United States ... with intent to defame the said government ... or to bring [it] ... into contempt or disrepute.”¹¹ Many citizens at the time – among them James Madison, a key delegate at the 1787 Constitutional Convention and the principal author of the Bill of Rights – argued that the Sedition Act was unconstitutional, either as a violation of the Free Speech Clause of the First Amendment or as an encroachment on the powers of the state governments (see Urofsky & Finkelman 2011, 204; Ellis 2001, 198-201). Other members of the Founding generation, including Alexander Hamilton and, quite obviously, a majority of the Congress that passed the Sedition Act, disagreed (though Hamilton himself did so halfheartedly) (Ellis 2001, 191). What then was the “original meaning” of the Free Speech Clause with respect to “defamation” of the government, or the original meaning of the Tenth Amendment, which “reserved to the States” all powers not “delegated” by the Constitution to the federal government? The answer depends on whether one credits the stated views of the (mostly Federalist) supporters of the Sedition Act or those of its (mostly Republican) opponents. Here also, it is far too simplistic to assign a single “original meaning” to these provisions, which by all evidence were contested from the moment of their origin.

Finally, consider the practical difficulties facing the judge seeking to identify original meaning even where it exists. The judge must, first, locate relevant historical evidence regarding what the appropriate collection of people thought the text meant (and in so doing must decide what evidence is relevant, which collection of people is appropriate, and what understandings or beliefs or other mental states of those people matter). She must then determine whether some of the relevant mental states of some of the people in question are in conflict and, if so, how to resolve that conflict. And she must, finally, figure out how to apply those historical mental states to the potentially very different and unforeseen facts as they exist today. As Justice Scalia (1989a, 856-57) himself puts it:

[I]t is often exceedingly difficult to plumb the original understanding of an ancient text. Properly done, the task requires the consideration of an enormous mass of material Even beyond that, it requires an evaluation of the reliability of that material And further still, it requires immersing oneself in the political and intellectual atmosphere of the time – somehow placing out of mind knowledge that we have which an earlier age did not, and putting on beliefs, attitudes, philosophies, prejudices and loyalties that are not those of our day. It is, in short, a task sometimes better suited to the historian than the lawyer.

All of which raises the obvious question: are we really constraining judges – who typically are not, after all, historians – in any meaningful way by asking them to search for original meaning in deciding contemporary cases? Many originalist theorists have been forced to admit that the answer is no. Randy Barnett and Lawrence Solum, two leading contemporary originalists, thus endorse what they call “constitutional construction” – essentially, discretionary judicial lawmaking – as a

¹¹ See Urofsky & Finkelman (2011, 201-05).

means of filling the many gaps in originalist interpretation (Barnett 2004, 118-30; Solum 2010). In Barnett's words (2004, 121):

For better or worse, ... the U.S. Constitution requires more than originalist interpretation to be applied to cases and controversies. Owing to the vagueness of language and the limits of historical inquiry, originalist interpretation may not result in a unique rule of law to be applied to a particular case or controversy.

Original meaning, like the text, is endemically underdeterminate. And while the problem may be most egregious in the context of constitutional provisions, which typically "use[] general concepts and abstract principles in place of specific rules" (Barnett 2004, 120), it will appear to some extent in the interpretation of any legal text.

4 "The Principal Weakness of the System": Participatory Adjudication and Judicial Constraint

The assumption that formalist techniques can meaningfully constrain judges is therefore deeply problematic. In this and the following two sections, I argue that this is not a great cause for concern – because the assumption that formalism is *necessary* to meaningfully constrain judges also is flawed. American judges are not constrained by the formalistic interpretation of legal rules. But they are constrained by the internal dynamics of the adversary system of adjudication; and they are constrained by the need to keep faith with substantive legal principles, even when those principles are not reducible to formalistic rules. I discuss the former source of constraint in this section and the latter in section 5.¹²

In the United States, as in most legal systems influenced heavily by the British, adjudication follows an "adversary" model: primary responsibility for initiating court cases, framing the factual and legal issues, investigating and proving the relevant facts, advancing the legal arguments, and shaping the remedies lies with the affected parties and their attorneys, not with the court itself. (I will focus the discussion here on civil disputes rather than criminal prosecutions, but most of the same party-driven dynamics apply in criminal cases – often to an even greater degree – although of course one of the parties to a criminal case is the government.)

In American civil adjudication:

- The disputing parties, not the court, choose whether and when to adjudicate a dispute. In the federal courts, moreover, disputes cannot be adjudicated unless the party seeking the court's intervention can show that it has suffered an "injury in fact," that its injury is "fairly traceable" to unlawful conduct by the defending party, and that a court order is "likely" to redress that injury. The mere existence of unlawful conduct, or of some general

¹² The discussion in this section is drawn from Peters (2011, 155-60, 310-28); see also Peters (1997, 347-60).

societal harm flowing from that conduct, is not enough to trigger adjudication.¹³

- The disputing parties (typically through their attorneys), not the court, determine which factual issues will be decided in resolving the dispute. The litigants investigate the relevant facts, assemble them in forms suitable for proof, and (within limits imposed by evidentiary rules) decide whether and how to present them to the trier of fact. Expert witnesses, where required, typically are retained and compensated by the litigants themselves, not by the court.
- The disputing parties also are chiefly or solely responsible for identifying the legal issues and developing the legal arguments relevant to the dispute. With few exceptions (most of which go to the court’s jurisdiction to hear the dispute in the first place), strong norms of judicial practice prevent judges from deciding cases based on legal grounds not argued by the parties.
- In many cases, the disputing parties have a substantial role in deciding who will make findings based on the proofs – whether it will be a judge or a jury, and in jury cases, who will sit on the jury.¹⁴ The jury option is itself a significant constraint on the judge’s authority that is not available in civil adjudication in most legal systems outside the United States.
- The disputing parties typically decide whether to settle their dispute (and thus terminate the adjudication) without need for court approval.¹⁵
- The disputing parties propose and argue the merits of remedies, such as compensatory damages or injunctive relief. While judges typically have substantial influence in the shaping of “specific” (injunctive) remedies, they generally cannot grant such remedies without being asked by the litigants to do so, and the form of specific remedies is determined by the proofs and arguments of the parties.

¹³ See, e.g., *Lujan v. Defenders of Wildlife*, 504 U.S. 555 (1992); *Allen v. Wright*, 468 U.S. 737 (1984).

¹⁴ The Seventh Amendment to the federal Constitution guarantees the right to a jury trial in many civil cases in federal court, but this right may be waived by the parties. In both federal and state courts, the litigants typically participate in the process of choosing jurors, known as *voir dire*, by asking questions of prospective jurors (either directly or by submitting them to the judge), by moving to exclude prospective jurors “for cause,” and by exercising “peremptory challenges” to strike a certain number of jurors without cause.

¹⁵ In fact, procedural rules and statutes in the United States increasingly encourage out-of-court settlement of disputes, and far more cases “settle” than go to trial in American courts (see Galanter & Cahill 1994). There are exceptions, however, to the baseline principle that judicial approval is not required for settlement, such as class-action lawsuits, where judicial approval is required in order to protect the interests of absent class members. See Federal Rule of Civil Procedure 23. And while the terms of settlements rarely require judicial approval, it is commonplace that judges use various means to encourage the litigants to settle (see Galanter & Cahill 1994).

- The disputing parties, not the trial judge or appellate judges, decide whether to appeal the trial court’s decision to a higher court and how to argue the appeal once it is taken. In the federal system, most appeals cannot be taken until there is a final judgment in the trial court, and only a party that is aggrieved by the trial court’s decision has standing to appeal it. No appellate court will review an issue – no matter how publicly important that issue may be – unless an aggrieved party seeks review of that issue.¹⁶
- Strong norms of judicial practice, and in some cases formal rules,¹⁷ require judges to issue written opinions explaining how their dispositive rulings are justified by the facts and the law and responding to the litigants’ arguments.

These prominent elements of litigant participation can and do exert meaningful constraint on the discretion of judges. American judges, unlike American legislators, cannot take action simply because they see a problem that needs solving. They must wait for that problem to generate a dispute between parties who stand to gain or lose something real and concrete from the resolution of that dispute. Judges then must act within the confines of that dispute as defined by the parties – deciding those legal and factual issues (and *only* those legal and factual issues) identified as important by the litigants, refraining from decision if the litigants choose to settle their dispute, and explaining and justifying their decision in terms that respond saliently to the litigants’ proofs and arguments. American judges are, in a very real sense, prisoners of the disputes they must resolve.¹⁸

This is true even in the context of important public-law litigation (which tends to be the focus of formalist concerns), although in that context the scope of the dispute, and thus of litigant participation, often is expanded. Consider the follow-

¹⁶ In the federal court system, the requirement of a final judgment as a condition of appeal is codified in 28 U.S.C. § 1291. There are a number of exceptions to this requirement, one of which allows the trial judge, of her own accord, to “certify” certain issues for immediate appeal. See 28 U.S.C. § 1292(b). The “final judgment” requirement for appeal is less stringent in many state court systems. In the federal courts, the requirement that an appeal from a final judgment may be taken only by an aggrieved party has been held to be at least partly constitutional in stature, flowing from Article III’s grant of the federal judicial power to decide only “cases” and “controversies.” See *Hollingsworth v. Perry*, 133 S. Ct. 2652 (2013); *United States v. Windsor*, 133 S. Ct. 2675, 2684-89 (2013).

¹⁷ Such as Federal Rule of Civil Procedure 52, which requires federal judges who serve as triers of fact to write opinions justifying their decisions on both factual and legal grounds.

¹⁸ Formalists often recognize the practical constraint imposed by the American adversary model of adjudication. Justice Scalia, for example, advocates strict adherence to justiciability requirements as a way to limit judicial power. See *United States v. Windsor*, 133 S. Ct. 2675, 2697, 2697-2705 (2013) (Scalia, J., dissenting) (criticizing the majority’s willingness to recognize standing to appeal in the case as reflecting “an exalted conception of the role of [the Supreme Court] in America”). This recognition sits somewhat uneasily alongside the professed belief that formalist interpretive techniques can meaningfully constrain judges. If formalism imposes significant constraint, it is unclear why the additional measure of strict justiciability limitations is necessary.

ing example taken from a significant and relatively recent Supreme Court adjudication.

In 2003, the Court decided the cases *Grutter v. Bollinger* and *Gratz v. Bollinger*.¹⁹ In *Grutter*, a white plaintiff who had been denied admission to the prestigious University of Michigan Law School sued the University to challenge, as a violation of the Fourteenth Amendment's Equal Protection Clause,²⁰ the Law School's "affirmative action" admissions policy, which gave a preference to members of certain racial minority groups. In *Gratz*, two white plaintiffs denied admission to the University's undergraduate program brought a similar challenge. By a 5-4 vote, the Court upheld the race-conscious Law School policy in *Grutter*; by a 6-3 vote, it struck down the crucially different undergraduate policy in *Gratz*.

One notable thing about the *Grutter* and *Gratz* decisions was their relative narrowness – a feature that was at least in part a function of the dispute-resolving constraints faced by the Court. In the lower courts, the litigants in each case had developed a substantial factual record regarding the particular details and effects of each affirmative-action policy at issue. The subtle differences between the Law School's policy and the undergraduate policy ended up determining the results: Justices Sandra Day O'Connor (who wrote the Court's opinion in *Grutter*) and Stephen Breyer voted to uphold the Law School's policy but to invalidate the undergraduate policy, on the ground that the former was "narrowly tailored" to achieve racial diversity while the latter was not. The decisions thus were neither broad endorsements nor broad condemnations of the use of race-conscious admissions in higher education. Instead they were relatively fact-specific rulings, capable of providing guidance to litigants and judges (and to universities seeking to avoid becoming litigants) in future similar cases, but not of conclusively resolving subsequent disputes with materially dissimilar facts.

Just how fact-specific the rulings were become evident a few years later, in *Parents Involved v. Seattle School District No. 1*, when the Court (after a crucial change in membership) invalidated race-conscious policies used to assign students to public schools.²¹ The litigants challenging these policies in *Parents Involved* were able to successfully distinguish that case from *Grutter*, convincing the Court that the particular policies in question, like that in the *Gratz* case, were not sufficiently "narrowly tailored" to serve the diversity objective. The reach of the *Grutter* and *Gratz* decisions thus was confined to the facts of the disputes that produced them, with subsequent litigants left free to argue for different results based on the differing facts of those subsequent disputes. This relatively narrow, com-

¹⁹ *Grutter v. Bollinger*, 539 U.S. 306 (2003); *Gratz v. Bollinger*, 539 U.S. 244 (2003).

²⁰ "No State shall ... deny to any person within its jurisdiction the equal protection of the laws."

²¹ See *Parents Involved in Community Schools v. Seattle School District No. 1*, 555 U.S. 701 (2007). Between 2003 (when *Grutter* and *Gratz* were decided) and 2007, Justice O'Connor had left the court, replaced by the more-conservative Justice Samuel Alito (who voted with the majority in *Parents Involved*). In addition, Chief Justice Rehnquist had died and been replaced by John Roberts, but that change did not affect the Court's attitude toward race-conscious affirmative action, which both Rehnquist and Roberts opposed.

mon-law style of case-by-case adjudication in effect makes each new set of litigants the authors, in part, of the legal rules that will end up binding them.

This is not to deny that relatively broad principles can be gleaned from *Grutter* and *Gratz*, as they can from most Supreme Court decisions. Those cases did establish as a general matter that the interest of racial diversity in higher education is sufficiently “compelling” to justify race-consciousness in admissions, provided the use of race is closely tailored to achieving that objective. But even this principle has proven vulnerable to the vagaries of particular disputes. In *Parents Involved*, for example, Chief Justice Roberts, writing for a plurality of four Justices, suggested that the diversity rationale might not apply in the context of elementary and secondary education (as opposed to higher education as in *Grutter*).²² The scope of the principles established by a decision is, to a great extent, in the hands of subsequent Courts, and thus is subject to the arguments of subsequent litigants.

The holdings of prior cases can be just as indeterminate as the text or the original understanding of the legal rules those cases interpret; the idea of absolute constraint by precedent is another formalist myth. The real constraining effect of precedent derives from the fact that litigants, in a common-law system like the American one, must use precedent as a basis for their legal arguments. Precedent affects the decision of which cases to litigate and of how to litigate them. And these factors in turn constrain the judges charged with deciding litigated cases. Once a case is properly before a court, judges will have substantial (though far from unlimited) discretion in interpreting the applicable precedents and other sources of legal norms. But which cases come before a court, and which precedents and other norms apply to them, are to a large extent outside the judges’ control.

The *Grutter* and *Gratz* decisions illustrate the influence of litigant participation, not just in the relative fact-specificity of their rulings, but in the overt responsiveness of those decisions to the participants’ arguments. In the trial courts, where factfinding occurs in the American system, the judges in these cases wrote lengthy opinions recounting in meticulous detail the parties’ arguments and proofs; the trial judge’s opinion in *Grutter* consumes fifty-two pages of the *Federal Supplement Second* reporter, twenty-three of which are devoted to a detailed, nonevaluative statement of the facts proved at trial.²³ At the Supreme Court level, the grounds of the Court’s decisions, and indeed of the opinions of the concurring and dissenting Justices, could be traced almost without exception to arguments made by one of the parties or by other litigants participating as *amici curiae*. The sole exception appeared in Justice John Paul Stevens’s dissent in *Gratz*, which opined that the plaintiffs in that case lacked standing to seek injunctive relief in the case, an ar-

²² See *Parents Involved*, 555 U.S. at 723 (“In upholding the admissions plan in *Grutter*, ... this Court relied upon considerations unique to institutions of higher education ...”).

²³ *Grutter v. Bollinger*, 127 F. Supp. 2d 821 (E.D. Mich. 2001). Federal district (trial) court judges are required by rule to expressly state the findings of fact and the rulings of law upon which their final judgments are based. See Federal Rule of Civil Procedure 52.

gument not made by the University or any of its allies.²⁴ This is an exception that proves the rule: ensuring that litigants have standing to sue – that they have something concrete to win or lose in a court case – is a way to preserve the dispute-resolving posture of American adjudication and thus to limit the scope of judicial authority.

None of this is to deny that American judges, particularly Supreme Court Justices, often retain considerable legal discretion to choose from multiple possible outcomes within the confines of a particular case, or that this choice often will be driven by extralegal factors. The point is simply that the dispute-resolving structure of American adjudication imposes real constraints on the content (and, just as importantly, the occasions) of judicial decisions. Strong procedural norms require American judges to decide in ways that are responsive, and that are seen to be responsive, to the proofs and arguments made by disputing parties. Substantial judicial constraint is an important practical effect of these norms.

5 “A ... Criterion ... Quite Separate from the Preferences of the Judge,” Part III: Judicial Constraint and Substantive Legal Principles

In addition to being practically constrained by the dynamics of dispute resolution, American judges are constrained by the obligation to enforce substantive legal principles, even when the otherwise-applicable legal rules are indeterminate. By “substantive legal principles,” I mean legal norms that lack the all-or-nothing quality of rules. Ronald Dworkin influentially distinguished between rules and principles in this way (Dworkin 1978, 22-28). Dworkin noted that rules “are applicable in an all-or-nothing fashion. If the facts a rule stipulates are given, then either the rule is valid, in which case the answer it supplies must be accepted, or it is not, in which case it contributes nothing to the decision” (id., 24). If a valid legal rule states “Vehicles are not allowed in the public park,” and we are given that Jones’s pickup truck is a “vehicle” and that Jones has driven his vehicle into the public park, then we know that Jones has violated the law.

Principles, in contrast, “stat[e] a reason that argues in one direction, but d[o] not necessitate a particular decision” (id., 26). Dworkin cites the example of the principle “No man shall profit from his own wrong” (id.). If we are given that Smith has profited from his own wrong – by, for example, publishing a best-selling book detailing his criminal exploits – we do not (by virtue of this) know for sure that Smith has violated the law. The principle in question “argues in [that] direction,” but it may, in any given case, be outweighed by competing principles (such as a principle limiting punishments for criminal acts to those expressly pro-

²⁴ *Gratz v. Bollinger*, 539 U.S. 244, 282 (2003) (Stevens, J., dissenting).

vided by the legislature) or by applicable rules (such as the guarantee of “the freedom of speech” in the Constitution’s First Amendment).

Dworkin also uses the concept of “legal principles” in a more specific sense, to identify non-rule-like norms flowing from “requirement[s] of justice or fairness or some other dimension of morality,” as opposed to “policies,” which are non-rule-like norms that “se[t] out a [social] goal to be reached” (id., 22). I want to put this distinction to one side, however, and use the terminology of “legal principles” in its more general sense, to refer to legal norms that are not reducible to all-or-nothing rules. “Principles” in this sense might include what Dworkin calls “policies.” In fact, principles in this sense might themselves be derived from all-or-nothing legal rules. And this possibility allows for the further possibility of judicial constraint in the following of rules, even when the rules themselves are indeterminate.²⁵

To illustrate what I mean, suppose a judge must decide whether the inclusion of an inoperative combat truck as part of a war memorial violates the “no vehicles in the public park” ordinance. Neither the ordinance’s text alone nor the “original meaning” of that text provides a clear resolution of the issue. But text and original meaning do not exhaust the possible sources of the law’s content.

Suppose that although the judge is unsure whether the combat truck is a “vehicle” within the meaning of the ordinance, she can readily conjure objects that clearly fall within Hart’s “core of settled meaning” of that word (see Hart 1958, 607): a working combat truck, for example, or a passenger car, or a motorcycle. Clearly the ordinance would apply to bar these objects from the park; and thus the judge can use these clear instances (we might call them *positive paradigms*) as clues to the *justification* of the ordinance – the best understanding of why the lawmaker created the ordinance and used certain language to express it. The question for the judge would be: *what it is about these quintessential “vehicles” that might explain the legislature’s decision to ban them from the park?* If there is some property or collection of properties that obvious vehicles possess and that justifies their exclusion from a public park, then the judge might be able to resolve the case at hand by asking whether the war-memorial truck also possesses that property or collection of properties (and thus also should be excluded from the park).²⁶

Suppose the judge notices a number of properties that these quintessential “vehicles” have in common: they are loud; they produce noxious fumes; they pose a danger to pedestrians in the park; they are made largely of metal; they carry one or more people from place to place; they must be driven or piloted by someone; they are relatively expensive. Some of these properties (the first three) might reasonably justify a statute banning “vehicles” from the public park. Parks are places peo-

²⁵ The discussion that follows is drawn from Peters (2011, 176-81).

²⁶ Cf. Hart & Sacks (1994, 1378): “Why would reasonable men, confronted with the law as it was, have enacted this new law to replace it? The most reliable guides to an answer will be found in the instances of unquestioned application of the statute.”

ple go, often with children, to play and relax; excessive noise, noxious fumes, and large speeding objects frequently will interfere with that function. On the other hand, it is difficult to construct a reasonable argument why other properties (the last four) would justify excluding vehicles from the park: nothing in particular about objects that are made of metal, or that carry people from place to place, or that require a driver, or that cost a lot of money is likely to interfere with the function of a public park. So the judge might hypothesize, based on these positive paradigms, that the justification of the “no vehicles” ordinance is to prevent excessive noise, noxious fumes, and danger to pedestrians in the park.

The judge might also notice that some objects that clearly are *not* vehicles – that is, cases that are *negative* paradigms – also possess one or more of these properties. Fireworks, for instance, produce lots of noise, noxious smoke, and danger to bystanders; and yet they clearly are not banned by the “no vehicles” ordinance. But this does not mean that the judge’s hypothesized justification of the ordinance (preventing noise, fumes, and danger) is wrong. Legislatures have limited time and resources, and often they will address only the most salient or usual manifestations of a problem (vehicles, which are common, but not fireworks, which are less so). Moreover, sometimes the reason for regulating one type of activity will, if applied to similar types of activity, be outweighed by countervailing reasons against regulation. (Banning fireworks from the park, for example, would ruin the city’s annual Independence Day celebration.) So the judge will be cautious about reading too much into the existence of a few negative paradigms. But the existence of a great many negative paradigms – lots of loud, smelly, dangerous things that clearly aren’t banned from the park – should cause the judge to rethink her hypothetical justification of the ordinance.

Relatedly, the judge is likely to consider the body of law apart from the ordinance in question, and to ask what understanding of the ordinance’s impact on that body of law makes the most sense. Suppose that the park independently is subject to a noise ordinance that prohibits sounds in excess of a certain decibel level. This would militate against an interpretation of the ordinance as justified in part by concerns about excessive noise; such an interpretation would render the ordinance partially redundant. Suppose, on the other hand, that the city allows organized soccer matches to be held in the park in front of loud groups of spectators; this too would militate against the “no excessive noise” justification of the ordinance, not because it would render the ordinance redundant, but rather because it would render it useless. A focus on the state of the law absent the ordinance also suggests a complementary inquiry into the particular problems or events that led the legislature to enact the ordinance. Perhaps the city council adopted the law immediately after a pedestrian was struck and killed by a car in the park; that would suggest that the ordinance is justified at least in part by the goal of preventing such accidents.

What the judge is looking for here – what I have called the “justification” of the ordinance – is a species of “principle” in the general sense in which that concept

was employed by Dworkin.²⁷ To conclude that the justification of the ordinance is to prevent excessive noise, noxious fumes, and danger to pedestrians in the park is not to conclude that there is a legal *rule* against these phenomena in the park. The applicable legal *rule*, after all, prohibits only one source of these harms (“vehicles”), not all instances of them. And as I’ve suggested, the limited scope of the rule might be explained by countervailing considerations, such as the independent benefits of allowing some non-vehicular sources of noise, fumes, and danger into the park. The justification of the ordinance is an example of Dworkin’s “principle” because it “argues in one direction” (Dworkin 1978, 26) – the direction of banning noisy, smelly, dangerous items from the park – but might be outweighed by countervailing principles or rules.

And yet the justification of a statute, or of any legal rule, differs from Dworkin’s freestanding “principles” because it is *attached* to a legal rule: it exists solely as a means of interpreting and applying that rule in cases where the rule otherwise is indeterminate. Banning loud, smelly, dangerous objects from the park might generally be a good idea, but the judge would not be authorized to do so based solely on the “principle” that it is a good idea; she may do so only as part of her task of interpreting and applying the legal rule that prohibits “vehicles” from the park. The judge’s authority for applying the principle that justifies the rule stems from her authority to apply the rule itself. In applying the justificatory principle, the judge just *is* applying the rule.

And this, finally, is the point: legal rules might be *applied* to particular cases in a meaningful sense – even when they are indeterminate in such cases according to formalist methods – through the use of *principles* (justifications, reasons, purposes, goals) that explain the existence of those rules and thus are embodied in them. Judges, then, can be constrained even by indeterminate legal rules – constrained to apply those rules in the way that best serves the justifications of them.

It is true that the search for justificatory principles behind legal rules unavoidably has a normative component. The judge looks, not just for anything that might explain a statute, but for a normatively attractive explanation, a *justification*. In the “no vehicles” case, for example, the judge will reject “preventing objects made of metal in the park” as a justification for the ordinance on purely normative grounds: that justification would be normatively unreasonable. Different judges may reach different conclusions regarding which potential justifications are normatively reasonable, or more normatively reasonable than others. So the search for justificatory principles is far from absolutely constraining; the possibility (many formalists would say the danger) that the case will turn on judicial discretion remains.

²⁷ Dworkin probably would refer to it more specifically as a “policy” – “a kind of standard that sets out a goal to be reached, generally an improvement in some economic, political, or social feature of the community” (Dworkin 1978, 22). Or it might be called the “purpose” behind the ordinance: “the set of reasons for making those words a fixed part of the body of the law” (Wellman 1987, 463).

Still, it is fair to say that the search for justification, if it is done in good faith, imposes meaningful though not complete constraint on the party doing the searching. To attempt to apply a legal rule consistently with its justification is a very different task than to engage in unfettered, all-things-considered moral reasoning. A judge might, in good faith, determine that the inoperative combat truck is not covered by the justification of the “no vehicles” ordinance and thus is not banned from the park, even if the judge herself finds the truck unsightly, or thinks it is morally offensive to celebrate war, and thus would have banned it from the park had she been the legislator. To apply a statute in a way consistent with its justification is to apply the statute, and to apply a statute is emphatically *not* to create law from whole cloth.

6 “A ... Criterion ... Quite Separate from the Preferences of the Judge,” Part IV: Judicial Constraint and Procedural Legal Principles

In the preceding two sections, I argued that American judges deciding cases are meaningfully constrained, despite the frequent failure of formalist techniques to provide determinate legal rules that they must follow. Judges are constrained by the dispute-resolving character of the adjudicative enterprise in which they are engaged: they can render decisions only in response to disputes that are initiated and argued by affected parties. And they are constrained by the need to interpret indeterminate rules in light of the principles that justify those rules.

Each of these sources of constraint, however, depends for its existence on judges’ willingness to honor a different sort of legal principle: not substantive legal principles but *procedural* ones. The foregoing discussion suggests the presence of three fundamental procedural principles capable of generating judicial constraint in American adjudication. They are principles of *impartiality*, of *responsiveness*, and of *faithfulness*.

Impartiality is the absence from a dispute-resolving procedure of extrinsic factors favoring one side of the dispute over another. Elsewhere I have explained at length why impartiality is a necessary ingredient of an acceptable dispute-resolving process (see Peters 2011, 78-81). Here it will suffice to state the fairly obvious proposition that the losing party to a dispute is unlikely to willingly accept the result if she believes the process that generated it was arbitrarily biased against her. The dispute-resolving structure of American adjudication thus presupposes a reasonably impartial judge or panel of judges.

This norm of judicial impartiality often finds expression in rules of procedure or judicial ethics, such as the statute requiring disqualification of federal judges “in any proceeding in which [their] impartiality might reasonably be ques-

tioned.”²⁸ But, as the open-ended normative language of this rule suggests, the norm itself is best thought of as a principle in the Dworkinian sense, albeit one that emanates chiefly or entirely from the constitutional requirement of “due process of law.” Nominees for federal judgeships, during Senate confirmation hearings, typically decline to state their views on active legal issues on the ground that the issue “might come before the court” should they be fortunate enough to serve on it. When they do so, it is not because some legal rule requires them to; the disqualification statutes and rules of judicial conduct have not been read so broadly and in any event do not apply to persons who are not yet judges.²⁹ Rather, it is because the nominees feel the strong pull of an unwritten principle of impartiality, one that requires them to avoid precommitment to particular resolutions of disputes.

This unwritten principle has real consequences for how judges decide cases. For a judge, to be impartial is, again, to avoid to the greatest extent possible the influence of extrinsic factors favoring one party over the other; and judicial predispositions and biases are such extrinsic factors. Judicial impartiality therefore requires what, in the words of the Supreme Court, “might be described as open-mindedness”:

This quality in a judge demands, not that he have no preconceptions on legal issues, but that he be willing to consider views that oppose his preconceptions, and remain open to persuasion, when the issues arise in a pending case. This sort of impartiality seeks to guarantee each litigant, not an *equal* chance to win the legal points in the case, but at least *some* chance of doing so.³⁰

Impartiality doesn’t guarantee that each litigant has “an *equal* chance to win,” because in most instances one litigant’s case, taking only the *relevant* or intrinsic factors into account, will be stronger than the other’s. But it does guarantee, so far as possible, that each litigant’s chance to win will not be distorted by irrelevant factors, including the judge’s own preexisting biases. The principle of judicial impartiality thus requires judges to take the proofs and arguments of both sides seriously, refraining from deciding the outcome until all the evidence is in. Somewhat more obviously, impartiality requires reasonable equity in the opportunity to present these proofs and arguments, preventing the judge from denying to one litigant the procedural opportunities granted to another. If honored in good faith, then, the principle of judicial impartiality imposes real constraint.

²⁸ 28 U.S.C. § 455(a).

²⁹ A ban on comment about active legal issues imposed on judicial nominees or candidates might run afoul of the Free Speech Clause of the First Amendment. The Supreme Court has invalidated on First Amendment grounds a state law forbidding an elective judicial candidate to “announce his or her views on disputed legal or political issues.” See *Republican Party of Minnesota v. White*, 536 U.S. 765 (2002).

³⁰ *Republican Party of Minnesota v. White*, 536 U.S. at 778 (emphases in original). The Court in *White* acknowledged that preserving judicial impartiality in this sense might be a compelling state interest, but it nonetheless invalidated the limitation on judicial-candidate speech in that case on the ground that it was not in fact intended to serve that interest.

Responsiveness is the requirement that judicial decisions be grounded in the proofs and arguments offered by the disputing litigants (see Peters 2001, 25-27). Lon Fuller described this requirement as one of “congruence” between the litigants’ efforts and the judge’s decision: “if this congruence is utterly absent ... then the adjudicative process has become a sham, for the parties’ participation in the decision has lost all meaning” (Fuller 1978, 388). For better or worse, the American litigant-driven model of adjudication reflects a more-general populist political commitment to bottom-up decision-procedures reflecting the views and interests of the affected persons, as opposed to top-down procedures embodying the supposed wisdom of professional experts (see Damaška 1986; Peters 1997, 349). American judges must be impartial, but they must be impartial in choosing from among competing proofs and arguments presented by the litigants themselves, not in choosing more generally from among all the possible proofs and arguments out there in the world, or from among those that would be made by ideally competent litigants.

This norm of responsiveness sometimes appears in procedural rules, like the federal-court rule requiring judges to “stat[e] on the record ... or ... in an opinion or memorandum” their factual findings and legal conclusions at the close of a bench trial.³¹ But it mostly operates, like the impartiality norm, as an unwritten background principle. American judges typically write opinions justifying their decisions, even if no legal rule requires them to; they typically explain in those opinions why they are choosing one litigant’s proofs or arguments over the other’s; and (as I mentioned in discussing the *Grutter* and *Gratz* decisions) they typically avoid deciding cases on grounds not argued by the litigants. The responsiveness principle, like the impartiality principle, thus imposes meaningful constraint on judges, for it limits the bases of their decision to those the litigants choose to put forward.

Finally, *faithfulness* is the requirement that a judge seek to conform her decision to the law, even where no legal rule provides a determinate resolution of the issue. In American adjudication, as presumably in other legal systems, virtually everyone agrees with the formalists that judges (barring extraordinary circumstances) must apply existing legal rules where those rules dictate determinate outcomes. Following Dworkin and others, however, I’ve suggested here that judicial faithfulness requires more than the mechanical implementation of determinate legal rules. It also requires an effort to decide in accordance with the principles that justify legal rules, even when the rules themselves are indeterminate.

That this is in fact how American judges are expected to behave has I think been demonstrated by Dworkin (see Dworkin 1978; Dworkin 1986). And that it is how American judges actually do behave, or at least present themselves (and, one

³¹ Federal Rule of Civil Procedure 52(a).

suspects, see themselves) as behaving,³² is evident from any judicial decision not clearly preordained by the text of a legal rule – which is to say, nearly any judicial decision, as cases with clearly preordained results rarely see the inside of a courtroom. The Supreme Court in the *Grutter* and *Gratz* cases did not throw up its collective hands in despair that the text of the legal rule at issue – the vague guarantee of “the equal protection of the laws” in the Constitution’s Fourteenth Amendment – did not determinately resolve the question at hand; nor did it simply declare that, absent a clear command of the text or of “original understanding,” its members would decide the issue by voting their own unfettered moral preferences. Instead, the Court decided the issue, or at least purported to decide it, in a way that was faithful to the Fourteenth Amendment without being clearly dictated by it. The fact that the members of the Court disagreed about what this duty of faithfulness required does not mean the duty did not exist or that they ignored that duty; it means only that they held differing views about how best to fulfill it.

Together these procedural principles – impartiality, responsiveness, faithfulness – can assert significant constraint on the discretion of American judges. That is, judges are constrained to the extent they honor these procedural principles. But what requires them to honor these principles?

Sometimes the answer is a written legal rule, like the one requiring judicial recusal in cases of questionable impartiality. Most of the time, though, the answer is much less concrete, and perhaps much less *legal*. It lies in expectations regarding the professional role of the judge, in longstanding practice, in legal training, in the peer pressure of lawyers and colleagues, in innate notions of justice and fairness. But these principles are no less real for lack of authoritative enactment as part of some legal code. After all, the requirements urged by legal formalists like Justice Scalia – attention to text, obeisance to “original meaning” – are themselves merely unwritten principles in this sense. A judge’s obligation to obey them ultimately depends on whether they work in practice – whether they serve the normative goals that judges believe themselves committed to.

My argument here has been that the worthy normative goals of democratic governance, and thus of judicial constraint in service of democracy, can and are served at least as well by adherence to the foundational principles of impartiality, responsiveness, and faithfulness as by reflexive allegiance to legal formalism.

³² There is of course much writing by American judges themselves regarding how they understand their duties. For a classic judicial statement of this understanding that accords with the principle of faithfulness, see Cardozo (1921).

References

- Aristotle (1941), "Nicomachean Ethics", in: Richard McKeon (ed.) (1941), *The Basic Works of Aristotle*, New York: Random House, 927-1112.
- Barnett, R. (2004), *Restoring the Lost Constitution: The Presumption of Liberty*, Princeton, New Jersey: Princeton University Press.
- Brest, P. (1980), "The Misconceived Quest for the Original Understanding," *Boston University Law Review* 60/1, 204-38.
- Cardozo, B. (1921), *The Nature of the Judicial Process*, New Haven, Connecticut: Yale University Press.
- Damaška, M. (1986), *The Faces of Justice and State Authority: A Comparative Approach to the Legal Process*, New Haven: Yale University Press.
- Dworkin, R. (1978), "The Model of Rules I", in *Taking Rights Seriously*, Cambridge, Massachusetts: Harvard University Press, 14-45.
- _____. (1985), "The Forum of Principle", in *A Matter of Principle*, Cambridge, Massachusetts: Harvard University Press, 33-71.
- _____. (1986), *Law's Empire*, Cambridge, Massachusetts: The Belknap Press of Harvard University Press.
- Ellis, J. (2001), *Founding Brothers: The Revolutionary Generation*, New York: Alfred A. Knopf.
- Fuller, L. (1958), "Positivism and Fidelity to Law – A Reply to Professor Hart", *Harvard Law Review* 71/4, 630-72.
- _____. (1978), "The Forms and Limits of Adjudication", *Harvard Law Review* 92/2, 353-409.
- Galanter, M., and M. Cahill (1994), "'Most Cases Settle': Judicial Promotion and Regulation of Settlement", *Stanford Law Review* 46/6, 1339-92.
- Hart, H. L. A. (1958), "Positivism and the Separation of Law and Morals", *Harvard Law Review* 71/4, 593-629.
- Hart, H. M., Jr., and A. Sacks (1994), *The Legal Process: Basic Problems in the Making and Application of Law*, edited by William N. Eskridge, Jr. and Philip P. Frickey, Westbury, New York: Foundation Press.
- Hellman, D. (2014), "An Epistemic Defense of Precedent", in: C. J. Peters (ed.) (2014), *Precedent in the United States Supreme Court*, New York: Springer, 63-76.
- Peters, C. (1997), "Adjudication as Representation", *Columbia Law Review* 97/2, 312-436.
- _____. (2001), "Persuasion: A Model of Majoritarianism as Adjudication", *Northwestern University Law Review* 96/1, 1-37.
- _____. (2011), *A Matter of Dispute: Morality, Democracy, and Law*, New York: Oxford University Press.
- _____. (2014), "Originalism, Stare Decisis, and Constitutional Authority", in: C. J. Peters (ed.) (2014), *Precedent in the United States Supreme Court*, New York: Springer, 189-226.
- Scalia, A. (1989a), "Originalism: The Lesser Evil", *University of Cincinnati Law Review* 57/3, 849-65.
- _____. (1989b), "The Rule of Law as a Law of Rules", *University of Chicago Law Review* 56/4, 1175-88.
- _____. (1997), "Common-Law Courts in a Civil-Law System: The Role of United States Federal Courts in Interpreting the Constitution and Laws", in: A. Gutmann (ed.) (1997), *A Matter of Interpretation: Federal Courts and the Law*, Princeton: Princeton University Press, 3-47.
- Solum, L. (2011), "What Is Originalism? The Evolution of Contemporary Originalist Theory", available at <http://ssrn.com/abstract=1825543>.
- _____. (2010), "The Interpretation-Construction Distinction", *Constitutional Commentary* 27/1, 95-118.
- Sunstein, C. (1997), "Justice Scalia's Democratic Formalism", *Yale Law Journal* 107/2, 529-68.
- Urofsky, M., and P. Finkelman (2011), *A March of Liberty: A Constitutional History of the United States, Volume I, From the Founding to 1900*, New York: Oxford University Press.

Wellman, V. (1987), "Dworkin and the Legal Process Tradition: The Legacy of Hart and Sacks",
Arizona Law Review 29/3, 413-474.