Fidelity to Text and Principle

By Jack M. Balkin

July 2007

The American Constitution Society takes no position on particular legal or policy initiatives. All expressions of opinion are those of the author or authors. ACS encourages its members to express their views and make their voices heard in order to further a rigorous discussion of important issues.
Fidelity to Text and Principle

Jack M. Balkin

Is our Constitution a living document that adapts to changing circumstances or must we interpret it according to its original meaning? For years now people have debated constitutional interpretation in these terms. But the choice is a false one. Constitutional interpretation requires fidelity to the original meaning of the Constitution and to the principles that underlie the text. The task of interpretation is to decide how best to apply them in current circumstances. This is the method of text and principle. It is faithful to the original meaning of the constitutional text and its underlying purposes. It is also consistent with a basic law whose reach and application evolve over time, a basic law that leaves to each generation the task of how to make sense of the Constitution’s words and principles. Although the constitutional text and principles do not change without subsequent amendment, their application and implementation can. That is the best way to understand the interpretive practices of our constitutional tradition and the work of the many political and social movements that have transformed our understandings of the Constitution’s guarantees.

I. Original Meaning versus Original Expected Application

Constitutional interpretations are not limited to applications specifically intended or expected by the framers and adopters of the constitutional text. For example, the Eighth Amendment’s prohibitions on “cruel and unusual punishments” bans punishments that are cruel and unusual as judged by contemporary application of these concepts (and underlying principles), not by how people living in 1791 would have applied those concepts and principles.

This marks the major difference between my approach and the one popularized by one of originalism’s most prominent champions, Justice Antonin Scalia.1 Justice Scalia agrees that we should interpret the Constitution according to “the original meaning of the text, not what the original draftsmen intended.”2 He also agrees that the original meaning of the text should be read in light of its underlying principles. But he insists that the concepts and principles underlying those words must be applied in the same way that they would have been applied when they were adopted. As he puts it, the principle underlying the Eighth Amendment “is not a moral principle of ‘cruelty’ that philosophers can play with in the future, but rather the existing society’s assessment of what is cruel. It means not . . . ‘whatever may be considered cruel from one generation to the next,’ but ‘what we consider cruel today [i.e., in 1791]’; otherwise it would be no protection against the moral perceptions of a future, more brutal generation. It is, in other words, rooted in the moral perceptions of the time.”3

---

2 See A Matter of Interpretation, supra note 2, at 38.
3 See id. at 145 (emphasis in original).
Scalia’s version of “original meaning” is not original meaning in my sense, but actually a more limited interpretive principle, original expected application. Original expected application asks how people living at the time the text was adopted would have expected it would be applied using language in its ordinary sense (along with any legal terms of art). Justice Scalia can accommodate new phenomena and new technologies—like television or radio—by analogical extension with phenomena and technologies that existed at the time of adoption. But this does not mean, Scalia insists, that “the very acts that were perfectly constitutional in 1791 (political patronage in government contracting and employment, for example) may be unconstitutional today.”

II. Mistakes and Achievements

Scalia realizes that his approach would allow many politically unacceptable results, including punishments that would clearly shock the conscience today. So he frequently allows deviations from his interpretive principles, making him what he calls a “faint-hearted originalist.” For example, Scalia accepts the New Deal settlement that gave the federal government vast powers to regulate the economy that most people in 1787 would never have dreamed of and would probably have strongly rejected.

Scalia’s originalism must be “faint-hearted” precisely because he has chosen an unrealistic and impractical principle of interpretation, which he must repeatedly leaven with respect for stare decisis and other prudential considerations. The basic problem with looking to original expected application for guidance is that it is inconsistent with so much of our existing constitutional traditions. Many federal laws securing the environment, protecting workers and consumers—even central aspects of Social Security—go beyond original expectations about federal power, not to mention independent federal agencies like the Federal Reserve Board and the Federal Communications Commission, and federal civil rights laws that protect women and the disabled from private discrimination. Even the federal government’s power to make paper money legal tender probably violates the expectations of the founding generation. The original expected application is also inconsistent with constitutional guarantees of sex equality for married women, with constitutional protection of interracial marriage, with the constitutional right to use contraceptives, and with the modern scope of free speech rights under the First Amendment.

---

4 See id. at 140-41 (emphasis in original).
5 See Scalia, Originalism, supra note 2, at 861-64.
6 See, e.g., Gonzales v. Raich, 545 U.S. 1, 33 (2005) (Scalia, J., concurring).
7 See Kenneth Dam, The Legal Tender Cases, 1981 SUP. CT. REV. 367, 389 (“difficult to escape the conclusion that the Framers intended to prohibit” use of paper money as legal tender); Nomination of Robert H. Bork to be Associate Justice of the Supreme Court of the United States: Hearings Before the S. Comm. on the Judiciary, 100th Cong. 84-85 (1987).
The standard response to this difficulty is that courts should retain nonoriginalist precedents (i.e., those inconsistent with original expectation) if those precedents are well established, if they promote stability, and if people have justifiably come to rely on them. Interpretive mistakes, even though constitutionally illegitimate when first made, can become acceptable because we respect precedent. As Scalia explains, “[t]he whole function of the doctrine” of stare decisis “is to make us say that what is false under proper analysis must nonetheless be held true, all in the interests of stability.”

There are four major problems with this solution. First, it undercuts the claim that legitimacy comes from adhering to the original meaning of the text adopted by the framers, and that decisions inconsistent with the original expected application are illegitimate. It suggests that legitimacy can come from public acceptance of the Supreme Court's decisions, or from considerations of stability or economic cost.

Second, under this approach, not all of the “incorrect” precedents receive equal deference. Judges will inevitably pick and choose which decisions they will retain and which they will discard based on pragmatic judgments about when reliance is real, substantial, justified or otherwise appropriate. These characterizations run together considerations of stability and potential economic expense with considerations of political acceptability—which decisions would be too embarrassing now to discard—and political preference—which decisions particularly rankle the jurist’s sensibilities. Thus, one might argue that it is too late to deny Congress’s power under the Commerce Clause to pass the Civil Rights Act of 1964 but express doubts about the Endangered Species Act. One might accept that states may not engage in sex discrimination but vigorously oppose the constitutional right to abortion or the unconstitutionality of anti-sodomy statutes. This play in the joints allows expectations-based originalism to track particular political agendas and allows judges to impose their political ideology on the law—they very thing that the methodology purports to avoid.

Third, allowing deviations from original expected application out of respect for precedent does not explain why we should not read these mistakes as narrowly as possible to avoid compounding the error, with the idea of gradually weakening and overturning them so as to return to more legitimate decisionmaking. If the sex equality decisions of the 1970's were mistakes, courts should try to distinguish them in every subsequent case with the goal of eventually ridding us of the blunder of recognizing equal constitutional rights for women.

This brings us to the final, and more basic problem: Our political tradition does not regard decisions that have secured equal rights for women, greater freedom of speech, federal power to protect the environment, and federal power to pass civil rights laws as mistakes that we must unhappily retain; it regards them as genuine achievements of American constitutionalism and sources of pride. These decisions are part of why we understand ourselves to be a nation that has grown freer and more democratic over time. No interpretive theory that regards equal constitutional rights for women as an unfortunate blunder that we are now simply stuck with

---

regulation); see also A MATTER OF INTERPRETATION, supra note 2, at 138 (contemporary First Amendment protections are “irreversible” “whether or not they were constitutionally required as an original matter”).

12 See A MATTER OF INTERPRETATION, supra note 2, at 139.
because of respect for precedent can be adequate to our history as a people. It confuses achievements with mistakes, and it maintains them out of a grudging acceptance. Indeed, those who argue for limiting constitutional interpretation to the original expected application are in some ways fortunate that previous judges rejected their theory of interpretation; this allows them to accept as a starting point nonoriginalist precedents that would now be far too embarrassing for them to disavow.

By contrast, a focus on text and principle views most, if not all of these achievements as plausible constructions of constitutional principles that underlie the constitutional text and that must be fleshed out in doctrine. Whatever the framers of the Fourteenth Amendment might have expected, equal rights for women are fully consistent with the original meaning of the Equal Protection Clause and the Privileges or Immunities Clause of the Fourteenth Amendment and its underlying principles of equal citizenship and opposition to class and caste legislation. We need not regard decisions recognizing women’s equal rights as mistakes: quite the contrary, they are our generation’s attempt to make sense of and implement the Constitution’s text and its underlying principles. These decisions—and others like them—do not sacrifice constitutional fidelity on the altar of precedent; they demonstrate how development of judicial doctrine over time can implement and maintain constitutional fidelity. It is rather those who would retreat from the achievements of our constitutional tradition or accept them only grudgingly who lack fidelity, because they lack faith in the ability and the authority of succeeding generations to accept the Constitution as their Constitution and to make constitutional text and constitutional principles their own.

Original expectation originalism cannot account for how social movements and post-enactment history shape our constitutional traditions. It holds that social movements and political mobilizations can change constitutional law through the amendment process of Article V. They can also pass new legislation, as long as that legislation does not violate original expected application—as much federal post-New Deal legislation might. But no matter how significant social movements like the civil rights movement and the women’s movement might have been in our nation’s history, no matter how much they may have changed Americans’ notion of what civil rights and civil liberties belong to them, they cannot legitimately alter the correct interpretation of the Constitution beyond the original expected application.

The model of text and principle views the work of social movements and post-enactment history quite differently. The constitutional text does not change without Article V amendment. But each generation of Americans can seek to persuade each other about how the text and its underlying principles should apply to their circumstances, their problems, and their grievances. And because conditions are always changing, new problems are always arising and new forms of social conflict and grievance are always being generated and discovered, the process of argument and persuasion about what the Constitution’s principles mean is never-ending.

When people try to persuade each other about how the Constitution and its principles apply to their circumstances, they naturally identify with the generation that framed the constitutional text and they claim that they are being true to its principles. They can and do draw analogies between the problems, grievances and injustices the adopters feared or faced and the problems, grievances, and injustices of our own day. They also can and do draw on the
experiences and interpretive glosses of previous generations—like the generation that produced the New Deal or the Civil Rights Movement—and argue that they are also following in their footsteps.

Most successful political and social movements in America’s history have claimed authority for change in just this way: either as a call to return to the enduring principles of the Constitution or as a call for fulfillment of those principles. Thus, the key tropes of constitutional interpretation by social movements and political parties are restoration on the one hand, and redemption on the other. Constitutional meaning changes by arguing about what we already believe, what we are already committed to, what we have promised ourselves as a people, what we must return to and what commitments remain to be fulfilled.

When political and social movements succeed in persuading other citizens that their interpretation is the right one, they replace an older set of implementing constructions and doctrines with a new one. These constructions and implementations may not be just or correct judged from the standpoint of later generations, and they can be challenged later on. But that is precisely the point. Each generation makes the Constitution their Constitution by calling upon its text and its principles and arguing about what they mean in their own time. Interpreting the Constitution’s text and principles is how each generation connects its values to the commitments of the past and carries forward the constitutional project of the American people into the future.

From the standpoint of text and principle, it matters greatly that there was a women’s movement in the 1960’s and 1970’s that convinced Americans that both married and single women were entitled to equal rights and that the best way to make sense of the Fourteenth Amendment’s principle of equal citizenship was to apply it to women as well as men, despite the original expected application of the adopters. The equal protection decisions of the 1970’s that gave heightened scrutiny to sex-based classifications are not “mistakes” that we must grudgingly live with. They are applications of text and principle that have become part of our constitutional tradition. They might be good or bad applications; they might be incorrect or incomplete. That is for later generations to judge. But when people accept them, as Americans accept the notion of equality for women today, they do not do so simply on the basis of reliance interests—i.e. that we gave women equal rights mistakenly in the 1970’s, and now it’s just too late to turn back. They do so in the belief that this is what the Constitution actually means, that this is the best, most faithful interpretation of constitutional text and principles.

Originalism based on original expected application fails because it cannot comprehend this feature of constitutional development except as a series of errors that it would now be too embarrassing to correct. Justice Scalia correctly notes that his reliance on nonoriginalist precedents is not consistent with originalism, but rather a “pragmatic exception.” And that is precisely the problem with his view: The work of social movements in our country’s history is not a “pragmatic exception” to fidelity to the Constitution. It is the lifeblood of fidelity to our Constitution—as an ongoing project of vindicating constitutional text and principle in history.

In this way, the theory of text and principle explains—in a way that original expectation originalism cannot—why the Constitution is more than the dead hand of the past, but is a

---

13See id. at 140.
continuing project that each generation takes on. It is not a series of orders from the past but a
vibrant conversation between generations, in which succeeding generations pledge faith in the
constitutional project of their forbearers and exercise fidelity to the Constitution by making the
Constitution their own.

None of this means that the original expected application is irrelevant or unimportant. It
helps us understand the original meaning of the text and the general principles that animated the
text. But it is important not as binding law but rather as an aid to interpretation, one among
many others. It does not control how we should apply the Constitution’s guarantees today,
especially as our world becomes increasingly distant from the expectations and assumptions of
the adopters’ era.

III. Implementing Text and Principles

Although the original expected application is not binding, the constitutional text certainly
is. That is because we have a written Constitution that is also enforceable law. We treat the
Constitution as law by viewing its text and the principles that underlie the text as legal rules and
legal principles. We look to the original meaning of the words because if the meaning of the
words changed over time, then the words will embrace different concepts than the people who
had the authority to create the text sought to refer to. We look to underlying principles because
when the text uses relatively abstract and general concepts, we must know which principles the
text presumes or is attempting to embrace. Sometimes the text refers to terms of art or uses
figurative or non-literal language, in that case we must try to figure out what principles
underlie that term of art or figurative or non-literal language. If we do not look to underlying
principles, then we may be engaged in a play on words and we will not be faithful to the
Constitution’s purposes. We can and should use history to discover the Constitution’s
underlying principles. But the principles we derive from history must be at roughly the same
level of abstraction as the text itself. The question is not what principles people specifically
intended but what principles the text enacts.

When the text is relatively rule-like, concrete and specific, underlying principles cannot
override the textual command. For example, the underlying goal of promoting maturity in a
President does not mean that we can dispense with the 35 year age requirement. But where the
text is abstract, general, or offers a standard, we must look to the principles that underlie the text
to make sense of the text and produce applications consistent with the text. Because the text
points to general and abstract concepts, these underlying principles will usually also be general
and abstract. Indeed, the fact that adopters chose text that features general and abstract concepts
is normally the best evidence that they sought to embody general and abstract principles of
constitutional law, whose scope, in turn, will have to be worked out by later generations.

---

14For example, the Copyright Clause in Article I, section 8 refers to “writings,” which is a non-literal use. It refers
to more than written marks on a page, but also includes printing and (probably) sculpture, motion pictures, and other
media of artistic and scientific communication. The term “due process of law” in the Fifth and Fourteenth
Amendment is a term of art; that is, it has a specialized legal meaning over and above the concatenation of the words
in the phrase. See id. at 37-38 (the text of the First Amendment must be construed as a synecdoche in which
“speech,” and “press” stand for a whole range of different forms of expression, including handwritten letters).
Some principles are directly connected to particular texts and help us understand how to apply those texts. Other principles we infer from the constitutional structure as a whole. For example, there is no single separation of powers clause in the Constitution; rather we must derive the principle of separation of powers from how the various institutions and structures outlined in the constitutional text relate to each other. The principle of democracy—which includes the subprinciple that courts should generally defer to majoritarian decisionmaking—is nowhere specifically mentioned in the constitutional text, and yet it may be the most frequently articulated principle in constitutional argument. It is, ironically, the principle that people most often use to object to courts inferring constitutional principles not specifically mentioned in the text. Although the principle of democracy does not directly appear in the text, we infer it from various textual features that presume democracy and from the basic character of our government as a representative and democratic republic.

Finally, many other materials gloss text and principles and help apply them to concrete circumstances. These include not only the original expected application but also post-enactment history, including the work of social movements that have changed our constitutional common sense, and judicial and non-judicial precedents. These materials offer a wide range of theories and interpretations about how to understand and apply the Constitution’s structures and guarantees. They are entitled to considerable weight. Precedents in particular not only implement and concretize principles, they also help settle difficult legal questions where reasonable people can and do disagree. Precedents also help promote stability and rule of law values. However, because glosses and precedents accumulate and change over time, and because they often point in contrasting directions, they are not always dispositive of constitutional meaning.

Constitutional doctrines created by courts and institutions and practices created by the political branches flesh out and implement the constitutional text and underlying principles. But they are not supposed to replace them. Doctrines, institutions and practices can implement the Constitution well or poorly depending on the circumstances, and some implementations that seem perfectly adequate at one point may come to seem quite inadequate or even perverse later on. Because the Constitution, and not interpretations of the Constitution, is the supreme law of the land, later generations may assert—and try to convince others—that the best interpretation of text and principle differs from previous implementing glosses, and hence that we should return to the best interpretation of text and principle, creating new implementing rules, practices and doctrines that will best achieve this end. The tradition of continuous arguments about how best to implement constitutional meaning generates changes in constitutional doctrines, practices, and law. That is why, ultimately, there is no conflict between fidelity to text and principle and practices of constitutionalism that evolve over time. Indeed, if each generation is to be faithful to the Constitution and adopt the Constitution’s text and principles as its own, it must take responsibility for interpreting and implementing the Constitution in its own era.

IV. Fidelity and Institutional Constraints

Expectations-based originalists may object that the text-and-principle approach is indeterminate when the text refers to abstract standards like “equal protection” rather than concrete rules. Therefore it does not sufficiently constrain judges. That might be so if text and
principle were all that judges consulted when they interpreted the Constitution. But in practice judges (and other constitutional interpreters) draw on a rich tradition of sources that guide and constrain interpretation, including pre and post-enactment history, original expected application, previous constitutional constructions and implementations, structural and inter-textual arguments, and judicial and non-judicial precedents. In practice, judges who look to text and principle face constraints much like those faced by judges who purport to rely on original expected application. As we have seen, the latter cannot and do not use original expected applications for a very large part of their work, because a very large part of modern doctrine is not consistent with original expected application. So even judges who claim to follow the original understanding are, in most cases, guided and constrained by essentially the same sources and modalities of argument as judges employing the method of text and principle.

I think there is a deeper problem with the objection that the method of text and principle does not sufficiently constrain judges. Many theories of constitutional interpretation conflate two different questions. The first is the question of what the Constitution means and how to be faithful to it. The second asks how a person in a particular institutional setting—like an unelected judge with life tenure—should interpret the Constitution and implement it through doctrinal constructions and applications. The first is the question of *fidelity*; the second is the question of *institutional responsibility*.

Theories about constitutional interpretation that conflate these two questions tend to view constitutional interpretation from the perspective of judges and the judicial role; they view constitutional interpretation as primarily a task of judges and they assess theories of interpretation largely in terms of how well they guide and limit judges. For example, one of the standard arguments for expectations-based originalism is that it will help constrain judges in a democracy. Alexander Bickel's theory of the passive virtues and Cass Sunstein's idea of minimalism, although often described as theories of constitutional interpretation, are actually theories about the judicial role and how judges should interpret the Constitution. So, too, obviously, are other theories of "judicial restraint." From the perspective of these theories, non-judicial interpreters are marginal or exceptional cases that we explain in terms of the standard case of judicial interpretation.

I reject this approach. Theories of constitutional interpretation should start with interpretation by citizens as the standard case; they should view interpretation by judges as a special case with special considerations created by the judicial role. In like fashion, constitutional interpretations by executive officials and members of legislatures are special cases that are structured by their particular institutional roles. Instead of viewing constitutional interpretation by citizens as parasitic on judicial interpretation, we should view it the other way around.

Why emphasize the citizen's perspective? Each generation must figure out what the Constitution's promises mean for themselves. Many of the most significant changes in constitutional understandings (e.g., the New Deal, the Civil Rights Movement, the second wave of American feminism) occurred through mobilizations and counter-mobilizations by social and political movements who offered competing interpretations of what the Constitution really means. Social and political movements argue that the way that Constitution has been interpreted
and implemented by judges or other political actors is wrong, and that we need to return to the Constitution's correct meaning and redeem the Constitution’s promises in our own day.

Often people do not make these claims in lawyerly ways, and usually they are not constrained by existing understandings and existing doctrine in the way that we want judges to be constrained. In fact, when social movements initially offer their constitutional claims, many people regard them as quite radical or “off the wall.” There was a time, for example, when the notion that the Constitution prohibited what we now call sex discrimination seemed quite absurd. Yet it is from these protestant interpretations of the Constitution that later constitutional doctrines emerge. Many of the proudest achievements of our constitutional tradition came from constitutional interpretations that were at one point regarded as crackpot and “off the wall.”

I hasten to add that most of these arguments go nowhere. Only a few have significantly changed how Americans look at the Constitution. Successful social and political movements must persuade other citizens that their views are correct, or, at the very least, they must convince people to compromise and modify their views. If movements are successful, they change the minds of the general public, politicians and courts. This influence eventually gets reflected in new laws, in new constitutional doctrines, and in new constitutional constructions. Successful social and political mobilization changes political culture, which changes constitutional culture, which, in turn, changes constitutional practices outside of the courts and constitutional doctrine within them.

The causal influences, of course, do not run in only one direction. Judicial interpretations like those in *Brown v. Board of Education*¹⁵ or *Miranda v. Arizona*¹⁶ can become important parts of our constitutional culture; they can be absorbed into ordinary citizens’ understandings of what the Constitution means, and they can act as focal points for citizen reaction. Nevertheless, we cannot understand how constitutional understandings change over time unless we recognize how social movements and political parties articulate new constitutional claims, create new constitutional regimes and influence judicial constructions.

To understand how these changes could be faithful to the Constitution, we must have a theory that makes the citizen’s perspective primary. I do not claim that all social mobilizations that produce changes in doctrine are equally legitimate or equally admirable. But some are both legitimate and admirable, and a theory of constitutional interpretation—which is also a theory of constitutional fidelity—must account for them. The text-and-principle approach can offer a much better explanation of how successful social and political movements make claims that are faithful to the Constitution than expectations-based originalism can. Indeed, expectations-based originalism is virtually useless for this purpose, because it views many of the most laudatory changes in our understandings of the Constitution as not faithful to the Constitution and therefore illegitimate. Thus, according to the logic of expectations-based originalism, the second wave of American feminism either cowed or bullied the federal judiciary into mistaken readings of the Fourteenth Amendment. Having blundered in the 1970’s, judges must now maintain these errors because of respect for precedent and political realities.

For similar reasons, expectations-based originalism cannot really constrain judges because too many present-day doctrines are simply inconsistent with it; as a result judges must pick and choose based on pragmatic justifications that are exceptions to the theory. Indeed, the exceptions threaten to swallow the theory in many areas of the law. Because expectations-based originalism conflates the question of constitutional fidelity with the question of judicial constraint, it offers the wrong answer to both questions.

Constraining judges in a democracy is important. But in practice most of that constraint does not come from theories of constitutional interpretation. It comes from institutional features of the political and legal system. Some of these are internal to law and legal culture, like the various sources and modalities of legal argument listed above. Others are “external” to legal reasoning but nevertheless strongly influence what judges produce as a group.

First, judges are subject to the same cultural influences as everyone else—they are socialized both as members of the public and as members of particular legal elites. Second, the system of judicial appointments and practices of partisan entrenchment determine and limit who gets to serve as a judge, helping ensure that most successful judicial candidates come from within the political and legal mainstream. Third, lower federal courts are bound to apply Supreme Court precedents. Fourth, the Supreme Court is a multi-member body whose decisions in contested cases are usually decided by the median or “swing” Justice. Over time, this keeps the Court’s work near the center of public opinion.

This combination of internal and external features constrains judicial interpretation in practice far more effectively than any single theory of interpretation ever could; it does much of the work in constructing which constitutional interpretations are reasonable and available to judges and which are “off the wall.” Equally important, this combination of internal and external factors keeps judicial decisions in touch with popular understandings of our Constitution’s basic commitments, continually translating, shaping and refining constitutional politics into constitutional law.

In short, we should not confuse the question of what it takes for actors in the system—including those actors who are not judges—to be faithful to the Constitution with the question of what features of the system constrain judicial interpretation. We must separate these questions to understand how constitutional fidelity occurs over time. When we do, we can also see why fidelity to original meaning and belief in a living Constitution are not at odds.

Fidelity to the Constitution means grappling with its text and its principles, applying them to our present circumstances, and making use of the entire tradition of opinions and precedents that have sought to vindicate and implement the Constitution. Reasonable people may disagree on what those principles mean and how they should apply. But the larger point about constitutional interpretation remains. We decide these questions by reference to text and principle, applying them to our own time and our own situation, and in this way making the Constitution our own. The conversation between past commitments and present generations is at the heart of constitutional interpretation. That is why we do not face a choice between living constitutionalism and fidelity to the original meaning of the text. The two are opposite sides of the same coin.