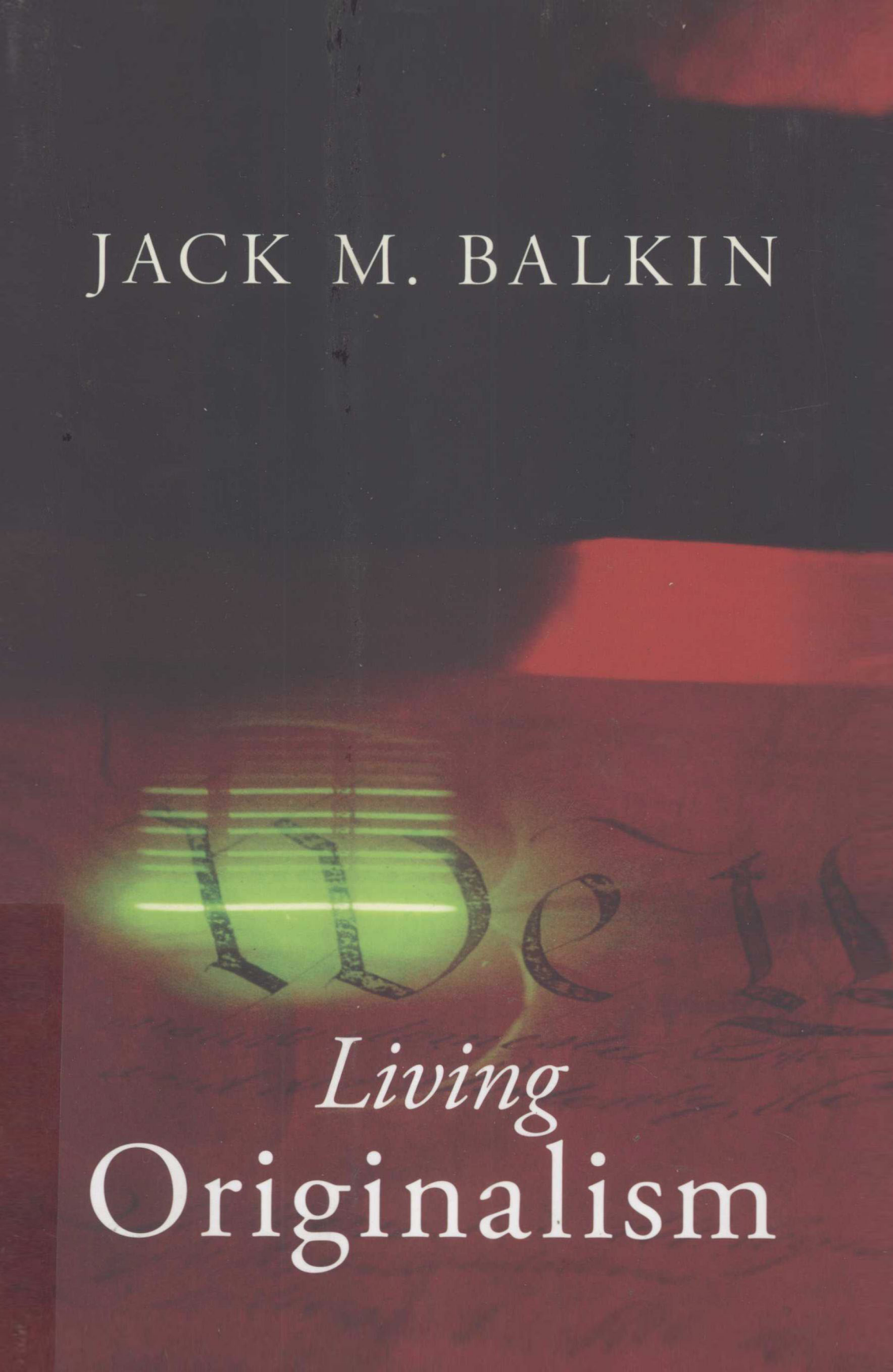
JACK M. BALKIN



Living Originalism

LIVING ORIGINALISM

Jack M. Balkin

The Belknap Press of Harvard University Press
Cambridge, Massachusetts
London, England
2011

For Reva Siegel and Robert Post

Library of Congress Cataloging-in-Publication Data

Balkin, Jack M., 1956–
Living originalism / Jack M. Balkin.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-674-06178-1 (alk. paper)
1. Constitutional law—United States—Philosophy. 2. Constitutional law—United States—Interpretation and construction. 3. Law—United States—Interpretation and construction 4. Origin (Philosophy) I. Title.
KF4552.B35 2011
342.73001—dc22 2011016124

CONTENTS

I

FIDELITY

II

CONSTRUCTION

III

CHANGE

I

FIDELITY

1

FIDELITY TO TEXT AND PRINCIPLE

Is our Constitution a living document that adapts to changing circumstances, or must we interpret it according to its original meaning? For many years people have debated constitutional interpretation in these terms. But the choice is a false one. Properly understood, these two views of the Constitution are compatible rather than opposed. Once we see why they are compatible, we will also understand how legitimate constitutional change occurs in the American constitutional system.

This book offers a constitutional theory, *framework originalism,* which views the Constitution as an initial framework for governance that sets politics in motion, and that Americans must fill out over time through constitutional construction. It also offers an associated theory of interpretation and construction, the method of *text and principle.* The method of text and principle requires fidelity to the original meaning of the Constitution, and in particular, to the rules, standards, and principles stated by the Constitution's text. It also requires us to ascertain and to be faithful to the principles that underlie the text, and to build out constitutional constructions that best apply the constitutional text and its associated principles in current circumstances.

The method of text and principle is both originalist and living constitutionalist. It is faithful to the original meaning of the constitutional text and to 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 implement the Constitution's words and principles. In each generation the American people are charged with the obligation to flesh out and implement text and principle in their own time. They do this through building political institutions, passing legislation, and creating precedents, both judicial and nonjudicial. These

constitutional constructions, in turn, shape how succeeding generations will understand and apply the Constitution in their time. 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.

The text of our Constitution is a framework. It is a basic plan for politics. The ratification of the Constitution begins a constitutional project that spans many generations.[1] Each generation must do its part to keep the plan going and to ensure that it remains adequate to the needs and the values of the American people. Americans fill the project out over time through constitutional politics. People contend with each other, trying to persuade each other about the best way to realize the constitutional plan and further its goals.

Keeping the plan going over time—especially given the many disagreements—requires faith in the constitutional project. This is a faith that the constitutional system as a whole is legitimate and worthy of our respect, or will come to be so over time, despite its many faults and imperfections. Thus, fidelity to the constitutional project—and to the Constitution itself—requires faith that the Constitution can and will eventually be redeemed. Fidelity to the Constitution requires that we believe that the project is worth continuing and struggling over, even if we also believe that many current interpretations are wrong or misguided.

Constitutional change is the product of this process of debate and striving over how to continue the plan. Americans try to persuade each other about the best meaning of constitutional text and principle in current circumstances. These debates and political struggles also help generate Americans' investment in the Constitution as their Constitution, even if they never officially consented to it; and they create a platform for the possibility—but not the certainty—of the Constitution's redemption in history.

What people call "constitutional interpretation" involves more than one activity. The first is the ascertainment of meaning. For example, in the First Amendment, does the word *speech* refer only to speaking, or does it point to a more general category of expression that might include writing, music, and painting? The second activity is constitutional construction—implementing and applying the Constitution using all of the various modalities of interpretation: arguments from history, structure, ethos, consequences, and precedent.[2] We might call the first activity "interpretation-as-ascertainment" and the second "interpretation-as-construction."

Much of what people call constitutional interpretation—especially by judges—is actually interpretation-as-construction. Judges build up systems of precedent that implement constitutional purposes and give the Constitution's guarantees and structures meaning in practice, and then they apply these systems of precedent to particular controversies. In the process they often create new doctrinal distinctions that will apply to future cases. The political branches also develop precedents over time through practices that flesh out the respective powers of the different branches of government. In this book I will often draw on the distinction between interpretation-as-ascertainment and interpretation-as-construction. When I speak of "interpretation" generally, I will include both types, as do most people.

Constitutional construction, however, involves far more than developing doctrines and precedents that implement the Constitution. All three branches of government build institutions and create laws and doctrines that serve constitutional purposes, that perform constitutional functions, or that reconfigure the relationships among the branches of the federal government, the states, and civil society. These activities build out the American state over time.

For example, Congress has created the various parts of the executive branch—like the Defense Department and the Justice Department—to help the president carry out his duties to faithfully execute the laws and perform other constitutional functions. As a result, the president is far more powerful today than anyone could have imagined in 1787. The Administrative Procedure Act of 1946 helps the courts review federal administrative actions for conformity to law. The Federal Reserve Act of 1913 and later amendments shifted the responsibility for monetary policy to an independent federal agency. The Civil Rights Act of 1964, the Voting Rights Act of 1965, and later civil rights measures created an enduring framework of national civil rights protection. The Social Security Act of 1935 and later social welfare laws constructed a social safety net administered by the federal government and the states. We might call these "state-building constructions." They build out the Constitution as they build out the country.

Not all of these state-building constructions look like interpretations of specific constitutional provisions, but they rely on assumptions about the Constitution's larger purposes and about what the Constitution permits or requires. Moreover, these constructions, once accepted in practice, may create durable expectations about what the Constitution means.

Courts often shape constitutional doctrines to make sense of these state-building exercises. In fact, the construction of the administrative state during the New Deal was the occasion for a major debate about constitutional interpretation that transformed how lawyers and judges understood the Constitution's guarantees. The New Deal and the civil rights revolution also changed how ordinary Americans understood the purposes and responsibilities of government and the rights of individual citizens.

To understand constitutional interpretation and the processes of constitutional change, we must pay as much attention to institutional development and state building as we do to judicial doctrines and decisions. To understand our Constitution, we must consider not only original meaning and judicial precedents, but also a wide variety of other state-building constructions that rely on interpretations of the Constitution and that provoke new interpretations.

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 First Amendment today does not protect only speech that people in 1791 would have protected from censorship. The Eighth Amendment's prohibitions on "cruel and unusual punishments" ban punishments that are cruel and unusual as judged by contemporary application of these concepts and principles, not by how people living in 1791 would have applied them.

The text of our Constitution contains different kinds of language. It contains determinate rules (the president must be thirty-five, there are two houses of Congress). It contains standards (no "unreasonable searches and seizures," a right to a "speedy" trial). And it contains principles (no prohibitions of the free exercise of religion, no abridgments of the freedom of speech, no denials of equal protection). If the text states a determinate rule, we must apply the rule because that is what the text offers us. If it states a standard, we must apply the standard. And if it states a general principle, we must apply the principle. Perhaps technically we should call this the method of "text, rule, standard, and principle," but "text and principle" is a far simpler shorthand.

The method of text and principle argues that we should pay careful attention to the reasons why constitutional designers choose particular kinds of language. Adopters use fixed rules because they want to limit discretion;

they use standards or principles because they want to channel politics through certain key concepts but delegate the details to future generations. When the Constitution uses vague standards or abstract principles, we must apply them to our own circumstances in our own time. When adopters use language that delegates constitutional construction to future generations, fidelity to the Constitution requires future generations to engage in constitutional construction. This is the essence of the method of text and principle.

This assumption marks the major difference between my approach and the one popularized by one of originalism's most prominent champions, Justice Antonin Scalia.[3] Justice Scalia agrees that we should interpret the Constitution according to "the original meaning of the text, not what the original draftsmen intended."[4] 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 formulated and applied in the same way that they would have been formulated and applied when they were adopted. As he puts it, the principle enacted in 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.*"[5]

Scalia's version of "original meaning" is not original meaning in my sense, but 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). Thus, the original expected application includes not only specific results, but also the way that the adopting generation would have expected the relevant constitutional principles to be articulated and applied.

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 *un*constitutional today."[6]

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."[7] 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.[8]

Scalia's originalism must be "faint-hearted" precisely because he has chosen an unrealistic and impractical principle of construction, which he must repeatedly leaven with respect for precedent and other prudential considerations. The basic problem with looking to original expected application for guidance is that 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. So too do 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.[9] The original expected application is also inconsistent with constitutional guarantees of sex equality for married women,[10] with constitutional protection of interracial marriage,[11] with the constitutional right to use contraceptives,[12] and with the modern scope of free-speech rights under the First Amendment.[13]

The standard response to this difficulty is that courts should retain "nonoriginalist" precedents (i.e., those inconsistent with original expected application) 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."[14]

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 antisodomy 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—the 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 1970s 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. If the New Deal settlement was thoroughly illegitimate, courts should find ways to strike down federal statutes, chip away at existing understandings, and ultimately overturn federal laws guaranteeing environmental quality, nondiscrimination, and workplace safety.

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 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.

An originalism that focuses on original expected applications cannot account for how political and social movements and post-enactment history shape our constitutional traditions. It cannot explain how succeeding generations build out the Constitution through constitutional construction. Original expectations originalism argues 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 political and 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 text and principle 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 emerging, the process of argument and persuasion about how to apply the Constitution's text and principles is never-ending.

When people try to persuade each other about how to interpret the Constitution, they naturally identify with the generation that framed the constitutional text and they claim that they are being true to its deepest principles. They can and do draw analogies between the problems, griev-

ances, 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 construction 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. In every generation, We the People of the United States make the Constitution our own by calling upon its text and its principles and arguing about what they mean in our own time. That 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 1960s and 1970s 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 laws and judicial decisions crafted in this period that ban sex discrimination 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—that we

mistakenly gave women equal rights in the 1970s 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 applications 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 his originalist commitments, but is rather a "pragmatic exception."[15] And that is precisely the problem with his view: The work of political and 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—it is how Americans vindicate their Constitution's text and principles in history.

None of this means that the original expected application is irrelevant or unimportant either to interpretation or construction. It helps us understand the original meaning of the text and the general principles that animated the text. For example, in Chapters 9 through 11, which apply the method of text and principle to the commerce clause and the Fourteenth Amendment, I draw heavily on the history of adoption to resolve ambiguities in original meaning and to suggest the best constructions. The point, however, is that original expected applications are not part of the text and they are not themselves binding law. Rather, like other aspects of pre- and post-ratification history, they are a method or modality of interpretation, one among many others. They do 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.

The Meaning of Original Meaning

The term *original meaning* can be confusing because we use the word *meaning* to refer to at least five different kinds of things: (1) semantic content ("What is the meaning of this word in English?");[16] (2) practical applications ("What does this mean in practice?"); (3) purposes or functions ("What is the meaning of life?"); (4) specific intentions ("I didn't mean to hurt you"); or (5) associations ("What does America mean to me?"). Thus, when we ask about the "meaning" of the equal protection clause, we could be asking (1) what concepts the words in the clause point to;